

Db 181 TWCTVLQNKQKVEFKIDIV 200

RESULT 28

US-08-328-500-9
; Sequence 9, Application US/08328500

; Patent No. 6673896

; GENERAL INFORMATION:

; APPLICANT: Maddon, Paul J.

; APPLICANT: Axel, Richard

; APPLICANT: Sweet, Richard W.

; APPLICANT: Arthos, James

; TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/328,500

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 0575/24577-CY

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 457 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-328-500-9

Query March 78.1%; Score 1029; DB 4; Length 457;

Best Local Similarity 100.0%; Pred. No. 3,4e-79;

Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLLVQLALPAATQGNKRVLGKGDVLELTCTASQKKSIOFHMNSNOIK 60

Db 1 MNRGVPRHLLLVQLALPAATQGNKRVLGKGDVLELTCTASQKKSIOFHMNSNOIK 60

QY 61 ILGNQGSFLTKGPKSNDRADSRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEEVOL 120

Db 61 ILGNQGSFLTKGPKSNDRADSRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEEVOL 120

QY 121 LVFGLTANSPTLHLLQGSGLTLTLSPSPGSPSVQCRSPRGNIOGKTLISVSOLELDSG 180

Db 121 LVFGLTANSPTLHLLQGSGLTLTLSPSPGSPSVQCRSPRGNIOGKTLISVSOLELDSG 180

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

APPLICANT: Littman, Dan R.

APPLICANT: Chess, Leonard

APPLICANT: Axel, Richard

APPLICANT: Weiss, Robin

APPLICANT: McDougall, J. S.

; TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN

; TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/466,368

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 24577-BI-B/JPM/AKC

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-278-0400

; TELEFAX: 212-391-0525

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 458 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-466-368-4

Query Match 78.1%; Score 1029; DB 3; Length 458;

Best Local Similarity 100.0%; Pred. No. 3,4e-79;

Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLLVQLALPAATQGNKRVLGKGDVLELTCTASQKKSIOFHMNSNOIK 60

Db 1 MNRGVPRHLLLVQLALPAATQGNKRVLGKGDVLELTCTASQKKSIOFHMNSNOIK 60

QY 61 ILGNQGSFLTKGPKSNDRADSRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEEVOL 120

Db 61 ILGNQGSFLTKGPKSNDRADSRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEEVOL 120

QY 121 LVFGLTANSPTLHLLQGSGLTLTLSPSPGSPSVQCRSPRGNIOGKTLISVSOLELDSG 180

Db 121 LVFGLTANSPTLHLLQGSGLTLTLSPSPGSPSVQCRSPRGNIOGKTLISVSOLELDSG 180

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

QY 181 TWCTVLQNKQKVEFKIDIV 200

Db 181 TWCTVLQNKQKVEFKIDIV 200

Chimeras

```
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,495
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/203,866
FILING DATE:
APPLICATION NUMBER: US/07/847,566
FILING DATE:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-417-495-5
```

```
Query Match 78.1%; Score 1029; DB 2; Length 462;
Best Local Similarity 100.0%; Pred. No. 3.4e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTQLALPPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHMKNNSNOIK 60
DB 1 MNRGVPRHLLVLTQLALPPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHMKNNSNOIK 60
QY 61 ILGNQGSFLLTKGPKSKLNDRADSRSLMDQGNFPLIKNLKIENSDTYICEVEDQKEVOL 120
DB 61 ILGNQGSFLLTKGPKSKLNDRADSRSLMDQGNFPLIKNLKIENSDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDTHLLQGSQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGSQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWICTVLQNOQKVEFKIDIV 200
DB 181 TWICTVLQNOQKVEFKIDIV 200
```

```
RESULT 31
US-08-284-391B-5
; Sequence 5, Application US/08284391B
; Patent No. 5851828
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Ebling LLP
; STREET: 176 Federal Street
; CITY: Boston
```

```
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Ebling, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-5
```

```
Query Match 78.1%; Score 1029; DB 2; Length 462;
Best Local Similarity 100.0%; Pred. No. 3.4e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTQLALPPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHMKNNSNOIK 60
DB 1 MNRGVPRHLLVLTQLALPPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHMKNNSNOIK 60
QY 61 ILGNQGSFLLTKGPKSKLNDRADSRSLMDQGNFPLIKNLKIENSDTYICEVEDQKEVOL 120
DB 61 ILGNQGSFLLTKGPKSKLNDRADSRSLMDQGNFPLIKNLKIENSDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDTHLLQGSQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGSQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWICTVLQNOQKVEFKIDIV 200
DB 181 TWICTVLQNOQKVEFKIDIV 200
```

```
RESULT 32
US-09-218-950-5
; Sequence 5, Application US/09218950
; Patent No. 6284240
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Ebling LLP
; STREET: 176 Federal Street
; CITY: Boston
```

```
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Pasteo for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218.950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284.391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195.395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847.566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665.961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35.238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-5

Query Match      78.1%; Score 1029; DB 3; Length 462;
Best Local Similarity 100.0%; Pred. No. 3.4e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALPAATQGNKVVGKGDVETLCTASQKSIQPHMKNNOIK 60
DB 1 MNRGVPRHLLVQLALPAATQGNKVVGKGDVETLCTASQKSIQPHMKNNOIK 60
QY 61 ILGNQGSFLTKGPKSKLNDRAISRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQGSFLTKGPKSKLNDRAISRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSDTHLLQGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWICTVLAQNKQKVEFKIDIV 200
DB 181 TWICTVLAQNKQKVEFKIDIV 200

RESULT 33
PCT-US92-01785-5
; Sequence 5, Application PC/TUS9201785
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
; TITLE OF INVENTION: Chimeras
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
```

```
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Mordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665.961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-01785-5

Query Match      78.1%; Score 1029; DB 5; Length 462;
Best Local Similarity 100.0%; Pred. No. 3.4e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALPAATQGNKVVGKGDVETLCTASQKSIQPHMKNNOIK 60
DB 1 MNRGVPRHLLVQLALPAATQGNKVVGKGDVETLCTASQKSIQPHMKNNOIK 60
QY 61 ILGNQGSFLTKGPKSKLNDRAISRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQGSFLTKGPKSKLNDRAISRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSDTHLLQGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWICTVLAQNKQKVEFKIDIV 200
DB 181 TWICTVLAQNKQKVEFKIDIV 200

RESULT 34
PCT-US95-00454-5
; Sequence 5, Application PC/TUS9500454
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
; TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Mordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
```

FILED DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/247001
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00454-5

Query Match 78.1%; Score 1029; DB 5; Length 462;
Best Local Similarity 100.0%; Pred. No. 3,4e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVYLQALLPAAQGNKVVLGKGGDTVEITCTASQKKSIOFHMKNNOIK 60
DB 1 MNRGVPRHLLVYLQALLPAAQGNKVVLGKGGDTVEITCTASQKKSIOFHMKNNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIETSDYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIETSDYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLQOGSLTLTLESPPGSSPSVQCRRSPKGNIOGKTLSSVQLELDPSG 180
DB 121 LVFGLTANSDTHLQOGSLTLTLESPPGSSPSVQCRRSPKGNIOGKTLSSVQLELDPSG 180
QY 181 TWCTCTVLONQKVEFKIDIV 200
DB 181 TWCTCTVLONQKVEFKIDIV 200

RESULT 35
US-08-417-495-6
Sequence 6, Application US/08417495
Patent No. 5843728
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Redirection of Cellular Immunity by
NUMBER OF INVENTION: Receptor
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,495
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/203,866
FILING DATE:
APPLICATION NUMBER: US/07/847,566
FILING DATE:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-417-495-6

Query Match 78.1%; Score 1029; DB 2; Length 532;
Best Local Similarity 100.0%; Pred. No. 4,1e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVYLQALLPAAQGNKVVLGKGGDTVEITCTASQKKSIOFHMKNNOIK 60
DB 1 MNRGVPRHLLVYLQALLPAAQGNKVVLGKGGDTVEITCTASQKKSIOFHMKNNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIETSDYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIETSDYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLQOGSLTLTLESPPGSSPSVQCRRSPKGNIOGKTLSSVQLELDPSG 180
DB 121 LVFGLTANSDTHLQOGSLTLTLESPPGSSPSVQCRRSPKGNIOGKTLSSVQLELDPSG 180
QY 181 TWCTCTVLONQKVEFKIDIV 200
DB 181 TWCTCTVLONQKVEFKIDIV 200

RESULT 36
US-08-284-391B-6
Sequence 6, Application US/08284391B
Patent No. 5851828
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
NUMBER OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elding LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Ebling, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-6

Query Match 78.1%; Score 1029; DB 2; Length 532;
Best Local Similarity 100.0%; Pred. No. 4.1e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVQLALPAAATQGNKVLGKGGDTVELCTASOKKSIOFHMNSNQIK 60
DB 1 MNRGVPFRHLVQLALPAAATQGNKVLGKGGDTVELCTASOKKSIOFHMNSNQIK 60
QY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120
DB 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120
QY 121 LVFGLTANSDTHLQGSGLTLTLESPPGSSPVQCSPRGKNIQGGKTLISVSQLELDQSG 180
DB 121 LVFGLTANSDTHLQGSGLTLTLESPPGSSPVQCSPRGKNIQGGKTLISVSQLELDQSG 180
QY 181 TWICTVLQONQKVEFKIDIV 200
DB 181 TWICTVLQONQKVEFKIDIV 200

RESULT 37
US-09-218-950-6
Sequence 6, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Ebling LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASCSEO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Ebling, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-6

Query Match 78.1%; Score 1029; DB 3; Length 532;
Best Local Similarity 100.0%; Pred. No. 4.1e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVQLALPAAATQGNKVLGKGGDTVELCTASOKKSIOFHMNSNQIK 60
DB 1 MNRGVPFRHLVQLALPAAATQGNKVLGKGGDTVELCTASOKKSIOFHMNSNQIK 60
QY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120
DB 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120
QY 121 LVFGLTANSDTHLQGSGLTLTLESPPGSSPVQCSPRGKNIQGGKTLISVSQLELDQSG 180
DB 121 LVFGLTANSDTHLQGSGLTLTLESPPGSSPVQCSPRGKNIQGGKTLISVSQLELDQSG 180
QY 181 TWICTVLQONQKVEFKIDIV 200
DB 181 TWICTVLQONQKVEFKIDIV 200

RESULT 38
PCT-US92-01785-6
Sequence 6, Application PC/TUS9201785
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991

ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 00786/119002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 532 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-01785-6

Query Match 78.1%; Score 1029; DB 5; Length 532;
 Best Local Similarity 100.0%; Pred. No. 4.1e-79;

Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALPPATQGNKVLGKKGDTVELTCTASQKKSIOFHMKNSQIK 60
 DB 1 MNRGVPRHLLVQLALPPATQGNKVLGKKGDTVELTCTASQKKSIOFHMKNSQIK 60
 QY 61 ILGNQSFLLTKGPKSLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120
 DB 61 ILGNQSFLLTKGPKSLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120
 QY 121 LVFGLTANSPTHLQGGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
 DB 121 LVFGLTANSPTHLQGGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
 QY 181 TWTCTVLQNKQKVEFKIDIV 200
 DB 181 TWTCTVLQNKQKVEFKIDIV 200

RESULT 39

PCT-US95-00454-6
 Sequence 6, Application PC/TUS9500454
 GENERAL INFORMATION:
 APPLICANT: Seed, Brian et al.
 TITLE OF INVENTION: Targeted Cycolysis of HIV-Infected
 TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 502 or 55SX
 OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 SOFTWARE: Wordperfect (Version 5.0)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/00454
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/847,566
 FILING DATE: March 6, 1992
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/665,961
 FILING DATE: March 7, 1991
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 00786/247001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 532 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-00454-6

Query Match 78.1%; Score 1029; DB 5; Length 532;
 Best Local Similarity 100.0%; Pred. No. 4.1e-79;

Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALPPATQGNKVLGKKGDTVELTCTASQKKSIOFHMKNSQIK 60
 DB 1 MNRGVPRHLLVQLALPPATQGNKVLGKKGDTVELTCTASQKKSIOFHMKNSQIK 60
 QY 61 ILGNQSFLLTKGPKSLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120
 DB 61 ILGNQSFLLTKGPKSLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120
 QY 121 LVFGLTANSPTHLQGGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
 DB 121 LVFGLTANSPTHLQGGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
 QY 181 TWTCTVLQNKQKVEFKIDIV 200
 DB 181 TWTCTVLQNKQKVEFKIDIV 200

RESULT 40

US-08-417-495-4
 Sequence 4, Application US/08417495
 Patent No. 5843728
 GENERAL INFORMATION:
 APPLICANT: Seed, Brian et al.
 TITLE OF INVENTION: Redirection of Cellular Immunity by
 TITLE OF INVENTION: Receptor
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 502 or 55SX
 OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 SOFTWARE: Wordperfect (Version 5.0)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/417,495
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/203,866
 FILING DATE:
 APPLICATION NUMBER: US/07/847,566
 FILING DATE:
 APPLICATION NUMBER: 07/665,961
 FILING DATE: March 7, 1991
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 00786/119002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-417-495-4

Query Match 78.1%; Score 1029; DB 2; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.6e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRVPRHLLVQLALLPAATGKRVVGGKDTVELCTASQKSIQFHMNSNQIK 60
DB 1 MNRVPRHLLVQLALLPAATGKRVVGGKDTVELCTASQKSIQFHMNSNQIK 60
QY 61 ILGQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKLKIEDSDTYICEVEDQKEEVQL 120
DB 61 ILGQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKLKIEDSDTYICEVEDQKEEVQL 120
QY 121 LVFGLTANSDTHLLQGSLTTLTSPGSSPSVQCRSPRGKNIQGKTLVSQLELDPSG 180
DB 121 LVFGLTANSDTHLLQGSLTTLTSPGSSPSVQCRSPRGKNIQGKTLVSQLELDPSG 180
QY 181 TWICTVLOKQKVEFKIDIV 200
DB 181 TWICTVLOKQKVEFKIDIV 200

RESULT 41

US-08-284-391B-4
Sequence 4, Application US/08284391B
Patent No. 5851828

GENERAL INFORMATION:

APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
CELLS BY CHIMERIC CD4 RECEPTOR - BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284.391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195.395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847.566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665.961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045

TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-4

Query Match 78.1%; Score 1029; DB 2; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.6e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRVPRHLLVQLALLPAATGKRVVGGKDTVELCTASQKSIQFHMNSNQIK 60
DB 1 MNRVPRHLLVQLALLPAATGKRVVGGKDTVELCTASQKSIQFHMNSNQIK 60
QY 61 ILGQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKLKIEDSDTYICEVEDQKEEVQL 120
DB 61 ILGQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKLKIEDSDTYICEVEDQKEEVQL 120
QY 121 LVFGLTANSDTHLLQGSLTTLTSPGSSPSVQCRSPRGKNIQGKTLVSQLELDPSG 180
DB 121 LVFGLTANSDTHLLQGSLTTLTSPGSSPSVQCRSPRGKNIQGKTLVSQLELDPSG 180
QY 181 TWICTVLOKQKVEFKIDIV 200
DB 181 TWICTVLOKQKVEFKIDIV 200

RESULT 42

US-09-218-950-4
Sequence 4, Application US/09218950
Patent No. 6284240

GENERAL INFORMATION:

APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
CELLS BY CHIMERIC CD4 RECEPTOR - BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218.950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284.391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195.395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847.566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665.961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-4

Query Match 78.1%; Score 1029; DB 3; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.6e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALLPAAQGNKVLGKGGDTVELTCTASOKKSIQFHKNSNQIK 60
DB 1 MNRGVPRHLLVQLALLPAAQGNKVLGKGGDTVELTCTASOKKSIQFHKNSNQIK 60
QY 61 ILGNQGSFLTKGPKSKNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPKSKNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDBTHLQGSQSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
DB 121 LVFGLTANSDBTHLQGSQSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
QY 181 TWCTVLOQOKKVEFKIDIV 200
DB 181 TWCTVLOQOKKVEFKIDIV 200

RESULT 43
PCT-US92-01785-4
Sequence 4, Application PC/TUS9201785
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
TITLE OF INVENTION: Chimeras
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear

MOLECULE TYPE: protein
PCT-US92-01785-4

Query Match 78.1%; Score 1029; DB 5; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.6e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALLPAAQGNKVLGKGGDTVELTCTASOKKSIQFHKNSNQIK 60
DB 1 MNRGVPRHLLVQLALLPAAQGNKVLGKGGDTVELTCTASOKKSIQFHKNSNQIK 60
QY 61 ILGNQGSFLTKGPKSKNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPKSKNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDBTHLQGSQSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
DB 121 LVFGLTANSDBTHLQGSQSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
QY 181 TWCTVLOQOKKVEFKIDIV 200
DB 181 TWCTVLOQOKKVEFKIDIV 200

RESULT 44
PCT-US95-00454-4
Sequence 4, Application PC/TUS9500454
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00454-4

Query Match 78.1%; Score 1029; DB 5; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.6e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLILVQLALPAAATQGNKVVLGKGGDTVELTCTASOKKSIQFHMKNNSQIK 60
DB 1 MNRGVPFRHLILVQLALPAAATQGNKVVLGKGGDTVELTCTASOKKSIQFHMKNNSQIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPDTHLLQGSGLTLTLSPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSPDTHLLQGSGLTLTLSPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCVTLQONOKKVEFKIDIV 200
DB 181 TWTCVTLQONOKKVEFKIDIV 200

RESULT 45
US-08-472-888A-6
Sequence 6, Application US/08472888A
Patent No. 6613746
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Walz, Gerd
TITLE OF INVENTION: AGP-ANTIBODY FUSION PROTEINS
TITLE OF INVENTION: AND RELATED MOLECULES AND METHODS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472.888A
CLASSIFICATION: 424
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/618,314
FILING DATE: 23-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/258001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 630 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-888A-6

Query Match 78.1%; Score 1029; DB 4; Length 630;
Best Local Similarity 100.0%; Pred. No. 5.2e-79;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLILVQLALPAAATQGNKVVLGKGGDTVELTCTASOKKSIQFHMKNNSQIK 60
DB 1 MNRGVPFRHLILVQLALPAAATQGNKVVLGKGGDTVELTCTASOKKSIQFHMKNNSQIK 60

DB 1 MNRGVPFRHLILVQLALPAAATQGNKVVLGKGGDTVELTCTASOKKSIQFHMKNNSQIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPDTHLLQGSGLTLTLSPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSPDTHLLQGSGLTLTLSPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCVTLQONOKKVEFKIDIV 200
DB 181 TWTCVTLQONOKKVEFKIDIV 200

RESULT 46
5223394-9
Patent No. 5223394
APPLICANT: WALLNER, BARBARA
TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
LINKAGE SIGNAL SEQUENCE
NUMBER OF SEQUENCES: 12
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/335,688
FILING DATE: 10-APR-1989
SEQ ID NO: 9
LENGTH: 295

Query Match 77.7%; Score 1023; DB 6; Length 295;
Best Local Similarity 99.5%; Pred. No. 6.1e-79;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLILVQLALPAAATQGNKVVLGKGGDTVELTCTASOKKSIQFHMKNNSQIK 60
DB 1 MNRGVPFRHLILVQLALPAAATQGNKVVLGKGGDTVELTCTASOKKSIQFHMKNNSQIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPDTHLLQGSGLTLTLSPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSPDTHLLQGSGLTLTLSPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCVTLQONOKKVEFKIDIV 200
DB 181 TWTCVTLQONOKKVEFKIDIV 200

RESULT 47
5223394-11
Patent No. 5223394
APPLICANT: WALLNER, BARBARA
TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
LINKAGE SIGNAL SEQUENCE
NUMBER OF SEQUENCES: 12
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/335,688
FILING DATE: 10-APR-1989
SEQ ID NO: 11
LENGTH: 318

Query Match 77.7%; Score 1023; DB 6; Length 318;
Best Local Similarity 99.5%; Pred. No. 6.7e-79;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLILVQLALPAAATQGNKVVLGKGGDTVELTCTASOKKSIQFHMKNNSQIK 60
DB 1 MNRGVPFRHLILVQLALPAAATQGNKVVLGKGGDTVELTCTASOKKSIQFHMKNNSQIK 60

QY 61 ILGNQGSFLTKGPKSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPKSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSVQLELODGS 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSVQLELODGS 180
QY 181 TWTCVLOQKQKVEFKIDIV 200
DB 181 TWTCVLOQKQKVEFKIDIV 200

RESULT 48
US-09-039-555B-15
; Sequence 15, Application US/09039555B
; Patent No. 6033856
; GENERAL INFORMATION:
; APPLICANT: Koerner, Kathrin
; APPLICANT: Mueller, Rolf
; APPLICANT: Sadiacek, Hans-Harald
; TITLE OF INVENTION: PROMOTER OF THE CDC25B GENE, ITS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,555B
; FILING DATE: 16-MAR-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 19710643.9
; FILING DATE: 14-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 016779/0131
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 458 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-039-555B-15

Query Match 77.7%; Score 1023; DB 3; Length 458;
Best Local Similarity 99.5%; Pred. No. 1.1e-78;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALPPATQGNKVVLGKGDVTELTCTASOKKSIQFHKNSNOIK 60
DB 1 MNRGVPRHLLVQLALPPATQGNKVVLGKGDVTELTCTASOKKSIQFHKNSNOIK 60
QY 61 ILGNQGSFLTKGPKSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPKSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120

QY 121 LVFGLTANSDTHLLOQOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSVQLELODGS 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSVQLELODGS 180
QY 181 TWTCVLOQKQKVEFKIDIV 200
DB 181 TWTCVLOQKQKVEFKIDIV 200

RESULT 49
US-09-517-605-3
; Sequence 3, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Litman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeek, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-517-605-3

Query Match 77.7%; Score 1023; DB 4; Length 458;
Best Local Similarity 99.5%; Pred. No. 1.1e-78;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALPPATQGNKVVLGKGDVTELTCTASOKKSIQFHKNSNOIK 60
DB 1 MNRGVPRHLLVQLALPPATQGNKVVLGKGDVTELTCTASOKKSIQFHKNSNOIK 60
QY 61 ILGNQGSFLTKGPKSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPKSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSVQLELODGS 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSVQLELODGS 180
QY 181 TWTCVLOQKQKVEFKIDIV 200
DB 181 TWTCVLOQKQKVEFKIDIV 200

RESULT 50
5223394-7
; Patent No. 5223394
; APPLICANT: WALLNER, BARBARA
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
; LINKAGE SIGNAL SEQUENCE
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/335,688
; FILING DATE: 10-APR-1989
; SEQ ID NO: 7
; LENGTH: 458
5223394-7

Query Match 77.1%; Score 1015; DB 6; Length 458;
Best Local Similarity 99.0%; Pred. No. 5.2e-78;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALPPATQGNKVVLGKGDVTELTCTASOKKSIQFHKNSNOIK 60
DB 1 MNRGVPRHLLVQLALPPATQGNKVVLGKGDVTELTCTASOKKSIQFHKNSNOIK 60

Db 1 MNRGVPFRHLLVQLALPAAATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIIRNLKIEDSDTYICEVEDQKEEVOL 120
Db 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIIRNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTLHLQGGSLTLTLSPGSSPSVQCSPRGKNIQGGKTLISVSQLELDQSG 180
Db 121 LVFGLTANSPTLHLQGGSLTLTLSPGSSPSVQCSPRGKNIQGGKTLISVSQLELDQSG 180
QY 181 TWCTCTVLQNKQKVEFKIDIV 200
Db 181 TWCTCTVLQNKQKVEFKIDIV 200

RESULT 51
US-08-466-368-2
; Sequence 2, Application US/08466368
; Patent No. 6093539
; GENERAL INFORMATION:
; APPLICANT: Maddon, Paul J.
; APPLICANT: Littman, Dan R.
; APPLICANT: Chess, Leonard
; APPLICANT: Axel, Richard
; APPLICANT: Weiss, Robin
; APPLICANT: McDougal, J. S.
; TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN
; TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,368
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 24577-E1-B/JPM/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-368-2

Query Match 76.8%; Score 1012; DB 3; Length 394;
Best Local Similarity 98.5%; Pred. No. 7.6e-78;
Matches 197; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALPAAATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60
Db 1 MNRGVPFRHLLVQLALPAAATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIIRNLKIEDSDTYICEVEDQKEEVOL 120
Db 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIIRNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTLHLQGGSLTLTLSPGSSPSVQCSPRGKNIQGGKTLISVSQLELDQSG 180
Db 121 LVFGLTANSPTLHLQGGSLTLTLSPGSSPSVQCSPRGKNIQGGKTLISVSQLELDQSG 180
QY 181 TWCTCTVLQNKQKVEFKIDIV 200
Db 181 TWCTCTVLQNKQKVEFKIDIV 200

Db 121 LVFGLTANSPTLHLQGGSLTLTLSPGSSPSVQCSPRGKNIQGGKTLISVSQLELDQSG 180
QY 181 TWCTCTVLQNKQKVEFKIDIV 200
Db 181 TWCTCTVLQNKQKVEFKIDIV 200

RESULT 52
US-08-328-500-2
; Sequence 2, Application US/08328500
; Patent No. 6673896
; GENERAL INFORMATION:
; APPLICANT: Maddon, Paul J.
; APPLICANT: Sweet, Richard
; APPLICANT: Athos, James
; TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,500
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/24577-CY
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-328-500-2

Query Match 76.8%; Score 1012; DB 4; Length 394;
Best Local Similarity 98.5%; Pred. No. 7.6e-78;
Matches 197; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALPAAATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60
Db 1 MNRGVPFRHLLVQLALPAAATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIIRNLKIEDSDTYICEVEDQKEEVOL 120
Db 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIIRNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTLHLQGGSLTLTLSPGSSPSVQCSPRGKNIQGGKTLISVSQLELDQSG 180
Db 121 LVFGLTANSPTLHLQGGSLTLTLSPGSSPSVQCSPRGKNIQGGKTLISVSQLELDQSG 180
QY 181 TWCTCTVLQNKQKVEFKIDIV 200
Db 181 TWCTCTVLQNKQKVEFKIDIV 200

RESULT 53
5223418-2

Patent No. 5223418
; APPLICANT: ARCURI, EDWARD J.; BRANNER, MARY E.; DONOVAN, MARY
; J.; GERBER, ROBERT G.; KELLER, JOHN A.
; TITLE OF INVENTION: METHOD OF IMPROVING THE YIELD OF
; HETEROLOGOUS PROTEINS PRODUCED BY STREPTOMYCES LIVIDANS
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/589,979
; FILING DATE: 28-SEP-1990
; SEQ ID NO: 2
; LENGTH: 394
5223418-2

Query Match 76.2%; Score 1003; DB 6; Length 394;
Best Local Similarity 98.0%; Pred. No. 4,4e-77;
Matches 196; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLIVLQALLPAATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNNOIK 60
DB 1 MNRGVPFRHLIVLQALLPAATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNNOIK 60
QY 61 ILGNQGSFLTKGPKKNDRADSRSLWDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPKKNDRADSRSLWDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTNSDTHLLOQGSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
DB 121 LVFGLTNSDTHLLOQGSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
QY 181 TWCTVLOKQKVEFKIDIV 200
DB 181 TWCTVLOKQKVEFKIDIV 200

RESULT 54
US-08-236-311-4
; Sequence 4, Application US/08236311
; Patent No. 5565335
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; APPLICANT: Gregory, Timothy J.
; TITLE OF INVENTION: Adheson Variants
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/236,311
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; APPLICATION NUMBER: 07/936190
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/842777
; FILING DATE: 18-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/250785
; FILING DATE: 28-SEP-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Haasek, Janet E.

REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 444PIC2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-236-311-4

Query Match 69.6%; Score 916; DB 1; Length 434;
Best Local Similarity 100.0%; Pred. No. 1.2e-69;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 QGNKVVLGKGGDTVELTCTASQKKSIOFHMKNNOIKILGNQGSFLTKGPKKNDRADSR 83
DB 56 QGNKVVLGKGGDTVELTCTASQKKSIOFHMKNNOIKILGNQGSFLTKGPKKNDRADSR 115
QY 84 RSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOLVFGLTNSDTHLLOQGSITLT 143
DB 116 RSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOLVFGLTNSDTHLLOQGSITLT 175
QY 144 ESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLOKQKVEFKIDIV 200
DB 176 ESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLOKQKVEFKIDIV 232

RESULT 55
US-08-457-918-4
; Sequence 4, Application US/08457918
; Patent No. 6117655
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; APPLICANT: Gregory, Timothy J.
; TITLE OF INVENTION: Adheson Variants
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,918
; FILING DATE: 1-JUN-1995
; CLASSIFICATION: 435
; APPLICATION NUMBER: 08/236311
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/936190
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/842777
; FILING DATE: 18-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/250785
; FILING DATE: 28-SEP-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Kudinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575

REFERENCE/DOCKET NUMBER: P0444P1C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8228
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 434 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-457-918-4

Query Match 69.6%; Score 916; DB 3; Length 434;
Best Local Similarity 100.0%; Pred. No. 1,2e-69;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 QGNKVLGKGDVLTCTASQKSIQFHMKNNOIKILNGSGFLTGP SKLNDRADSR 83
DB 56 QGNKVLGKGDVLTCTASQKSIQFHMKNNOIKILNGSGFLTGP SKLNDRADSR 115
QY 84 RSLMDQGNFPLIIKNLIKIEDSDTYICEVEDQKEEVQLVFGLTANSDTHLLOQGS LTTLES 143
DB 116 RSLMDQGNFPLIIKNLIKIEDSDTYICEVEDQKEEVQLVFGLTANSDTHLLOQGS LTTLES 175
QY 144 ESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLOKQKVEFKIDIV 200
DB 176 ESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLOKQKVEFKIDIV 232

RESULT 56
US-08-630-172-17
Sequence 17, Application US/08630172
Patent No. 6060054
GENERAL INFORMATION:
APPLICANT: Staerz, Uwe
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, 35th Floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630.172
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-36
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 410 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-630-172-17

Query Match 69.4%; Score 914; DB 3; Length 410;
Best Local Similarity 79.5%; Pred. No. 1,6e-69;
Matches 186; Conservative 9; Mismatches 27; Indels 12; Gaps 4;

QY 26 NKVVLGKGGDVELTCTASQKSIQFHMKNNOIKILNGSGFLTGP SKLNDRADSR 85
DB 1 NKVVLGKGGDVELTCTASQKSIQFHMKNNOIKILNGSGFLTGP SKLNDRADSR 60
QY 86 LMDQGNFPLIIKNLIKIEDSDTYICEVEDQKEEVQLVFGLTANSDTHLLOQGS LTTLES 145
DB 61 LMDQGNFPLIIKNLIKIEDSDTYICEVEDQKEEVQLVFGLTANSDTHLLOQGS LTTLES 120
QY 146 PRGSSPVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLOKQKVEFKIDIV----P 201
DB 121 PRGSSPVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLOKQKVEFKIDIVLAEP 180
QY 202 RASAL-PAPPTGSA LPPDPTASALPDPASALPALAVLSFLGLGVA CVL 254
DB 181 RGPITKCPCKCPAPN-----LLGSPSVFIFFPKIKV-LMISLSPIYTCV 227

RESULT 57
US-09-375-419-17
Sequence 17, Application US/09375419
Patent No. 6264950
GENERAL INFORMATION:
APPLICANT: Staerz, Uwe
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, 35th Floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/375.419
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630.172
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-36
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 410 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-375-419-17

Query Match 69.4%; Score 914; DB 3; Length 410;
Best Local Similarity 79.5%; Pred. No. 1,6e-69;
Matches 186; Conservative 9; Mismatches 27; Indels 12; Gaps 4;

QY 26 NKVVLGKGGDVELTCTASQKSIQFHMKNNOIKILNGSGFLTGP SKLNDRADSR 85
DB 1 NKVVLGKGGDVELTCTASQKSIQFHMKNNOIKILNGSGFLTGP SKLNDRADSR 60
QY 86 LMDQGNFPLIIKNLIKIEDSDTYICEVEDQKEEVQLVFGLTANSDTHLLOQGS LTTLES 145
DB 61 LMDQGNFPLIIKNLIKIEDSDTYICEVEDQKEEVQLVFGLTANSDTHLLOQGS LTTLES 120
QY 146 PRGSSPVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLOKQKVEFKIDIV----P 201

Db 121 PGGSSPVQCRSPRGKNIQGGKTLVSQLELDSDGTCTVLOKQKVEFKIDIVLAEP 180
Qy 202 RASAL-PAPPTGSALPDPTASALPDPPAASALPALAVISFLGLGAGVAVL 254
181 RGPTRKCPCKPCAPFN-----LLGGPSVFIFPPKIDV-LMTSLSPITCTV 227

RESULT 58
US-09-100-409A-1
Sequence 1, Application US/09100409A
Patent No. 6090388
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
TITLE OF INVENTION: IMMUNE DISORDERS
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100.409A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-751-6849
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-1

Query Match 68.3%; Score 899; DB 3; Length 433;
Best Local Similarity 100.0%; Pred. No. 3.2e-68;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 KVVLGKGGDTVELTCTASOKKSIQFHMKNKNOIKILGNQSGFLTKGPSKLNDRADSRSL 86
Db 2 KVVLGKGGDTVELTCTASOKKSIQFHMKNKNOIKILGNQSGFLTKGPSKLNDRADSRSL 61
Qy 87 WDOGNFPLIIKNIKIEDSDTYICEVEDQKEEVQLVFGILTANSDTHLLOQOSLTLTLESP 146
Db 62 WDOGNFPLIIKNIKIEDSDTYICEVEDQKEEVQLVFGILTANSDTHLLOQOSLTLTLESP 121
Qy 147 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDSDGTCTVLOKQKVEFKIDIV 200
Db 122 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDSDGTCTVLOKQKVEFKIDIV 175

RESULT 59
US-08-021-601-12
Sequence 12, Application US/08021601
Patent No. 5591631
GENERAL INFORMATION:
APPLICANT: Leppia, Stephen H.

APPLICANT: Klimpel, Kurt R.
APPLICANT: Nichols, Peter J.
APPLICANT: Atora, Naveen
APPLICANT: Singh, Yogendra
TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
TITLE OF INVENTION: RELATED METHODS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Needle & Rosenberg, P.C.
STREET: 133 Carnegie Way, Suite 400
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,601
FILING DATE: 19930212
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Spratt, Gwendolyn D.
REGISTRATION NUMBER: 36,016
REFERENCE/DOCKET NUMBER: 1414.057
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 903 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-021-601-12

Query Match 68.3%; Score 899; DB 1; Length 903;
Best Local Similarity 100.0%; Pred. No. 8.5e-68;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 KVVLGKGGDTVELTCTASOKKSIQFHMKNKNOIKILGNQSGFLTKGPSKLNDRADSRSL 86
Db 727 KVVLGKGGDTVELTCTASOKKSIQFHMKNKNOIKILGNQSGFLTKGPSKLNDRADSRSL 786
Qy 87 WDOGNFPLIIKNIKIEDSDTYICEVEDQKEEVQLVFGILTANSDTHLLOQOSLTLTLESP 146
Db 787 WDOGNFPLIIKNIKIEDSDTYICEVEDQKEEVQLVFGILTANSDTHLLOQOSLTLTLESP 846
Qy 147 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDSDGTCTVLOKQKVEFKIDIV 200
Db 847 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDSDGTCTVLOKQKVEFKIDIV 900

RESULT 60
US-08-082-849B-12
Sequence 12, Application US/08082849B
Patent No. 5677274
GENERAL INFORMATION:
APPLICANT: Leppia, Stephen H.
APPLICANT: Klimpel, Kurt R.
APPLICANT: Atora, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
TITLE OF INVENTION: Related Methods
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California

```

: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/082,849B
: FILING DATE: 25-JUN-1993
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/021,601
: FILING DATE: 12-FEB-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Weber, Kenneth A.
: REGISTRATION NUMBER: 31,677
: REFERENCE/DOCKET NUMBER: 15280-161-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 903 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-082-849B-12

Query Match      68.3%; Score 899; DB 1; Length 903;
Best Local Similarity 100.0%; Pred. No. 8.5e-68;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTQGPSKINDRADSRSL 86
DB 727 KVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTQGPSKINDRADSRSL 786

QY 87 WDOGNFPLIITKNIKIEDSDTYICEVEDQKEVOLVFGLTANSSTHLLQGOSLTLTLESP 146
DB 787 WDOGNFPLIITKNIKIEDSDTYICEVEDQKEVOLVFGLTANSSTHLLQGOSLTLTLESP 846

QY 147 PGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWTCTYLQNKQKVEFKIDIV 200
DB 847 PGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWTCTYLQNKQKVEFKIDIV 900

RESULT 61
PCT-US94-01624-12
: Sequence 12, Application PC/TUS9401624
: GENERAL INFORMATION:
: APPLICANT: Leppla, Stephen H.
: APPLICANT: Klimpel, Kurt R.
: APPLICANT: Arora, Naveen
: APPLICANT: Singh, Yogendra
: APPLICANT: Nicholas, Peter J.
: TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
: TITLE OF INVENTION: RELATED METHODS
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: TOWNSEND AND TOWNSEND KHOURI and CREW
: STREET: Steuart Street Tower, 20th floor, One Market
: STREET: Plaza
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94105
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/01624
```

```

: FILING DATE: June 25, 1993
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Weber, Kenneth A.
: REGISTRATION NUMBER: 31,677
: REFERENCE/DOCKET NUMBER: 15280-115
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 543-9600
: TELEFAX: (415) 543-5043
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 903 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US94-01624-12

Query Match      68.3%; Score 899; DB 5; Length 903;
Best Local Similarity 100.0%; Pred. No. 8.5e-68;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTQGPSKINDRADSRSL 86
DB 727 KVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTQGPSKINDRADSRSL 786

QY 87 WDOGNFPLIITKNIKIEDSDTYICEVEDQKEVOLVFGLTANSSTHLLQGOSLTLTLESP 146
DB 787 WDOGNFPLIITKNIKIEDSDTYICEVEDQKEVOLVFGLTANSSTHLLQGOSLTLTLESP 846

QY 147 PGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWTCTYLQNKQKVEFKIDIV 200
DB 847 PGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWTCTYLQNKQKVEFKIDIV 900
```

```

RESULT 62
US-08-867-149-1
: Sequence 1, Application US/08867149
: Patent No. 5912176
: GENERAL INFORMATION:
: APPLICANT: Wang, Chang Yi
: TITLE OF INVENTION: Antibody Against a Host Cell
: TITLE OF INVENTION: Antigen Complex for Pre- and Post-Exposure
: TITLE OF INVENTION: Protection from Infection by HIV Primary Isolates
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Maria C.H. Lin
: STREET: 345 Park Avenue
: CITY: New York
: STATE: NY
: COUNTRY: USA
: ZIP: 10154-0053
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version
: SOFTWARE: #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/867,149
: FILING DATE: 28-Feb-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/657,149
: FILING DATE: 03-June-1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Maria C.H. Lin
: REGISTRATION NUMBER: 29,323
: REFERENCE/DOCKET NUMBER: 1151-4145
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212)415-8745
: TELEFAX: (212)751-6849
: INFORMATION FOR SEQ ID NO: 1:
```

SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-867-149-1

Query Match 68.2%; Score 898; DB 2; Length 433;
Best Local Similarity 99.4%; Pred. No. 3,9e-68;
Matches 174; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 26 NKVLGKKGDTVELTCTASQKSIQFHMKNNOIKILGNQGSFLTKGPSKLNDRADSRSS 85
DB 1 NKVLGKKGDTVELTCTASQKSIQFHMKNNOIKILGNQGSFLTKGPSKLNDRADSRSS 60
QY 86 LMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDFHLLQGSLLTLES 145
DB 61 LMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDFHLLQGSLLTLES 120
QY 146 PGSSSPVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLONQKVEFKIDIV 200
DB 121 PGSSSPVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLONQKVEFKIDIV 175

RESULT 63
US-08-808-374-1
Sequence 1, Application US/08808374
Patent No. 5961976

GENERAL INFORMATION:
APPLICANT: Wang, Chang YI
TITLE OF INVENTION: Antibody Against a Host Cell
TITLE OF INVENTION: Antigen Complex for Pre- and Post-Exposure
TITLE OF INVENTION: Protection from Infection by HIV Primary Isolates
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,374
FILING DATE: 28-Feb-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/657,149
FILING DATE: 03-June-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4145
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-808-374-1

Query Match 68.2%; Score 898; DB 2; Length 433;
Best Local Similarity 99.4%; Pred. No. 3,9e-68;
Matches 174; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 26 NKVLGKKGDTVELTCTASQKSIQFHMKNNOIKILGNQGSFLTKGPSKLNDRADSRSS 85
DB 1 NKVLGKKGDTVELTCTASQKSIQFHMKNNOIKILGNQGSFLTKGPSKLNDRADSRSS 60
QY 86 LMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDFHLLQGSLLTLES 145
DB 61 LMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDFHLLQGSLLTLES 120
QY 146 PGSSSPVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLONQKVEFKIDIV 200
DB 121 PGSSSPVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLONQKVEFKIDIV 175

RESULT 64
5171838-13
Patent No. 5171838
APPLICANT: CHIBA, YUKIOBU
TITLE OF INVENTION: LEU3A BINDING PEPTIDES
NUMBER OF SEQUENCES: 24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/526,921
FILING DATE: 22-MAY-1990
SEQ ID NO:13:
LENGTH: 433
5171838-13

Query Match 54.5%; Score 718; DB 6; Length 433;
Best Local Similarity 83.4%; Pred. No. 6,9e-53;
Matches 151; Conservative 6; Mismatches 10; Indels 14; Gaps 4;

QY 27 KVLGKKGDTVELTCTASQKSIQFHMKNNOIKILGNQGSFLTKGPSKLNDRADSRSS- 85
DB 2 KVLGKKGDTVELTCTASQKSIQFHMKNNOIKILGNQGSFLTKGPSKLNDRADSRSS 61
QY 86 ----LMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDFHLLQGSLL 139
DB 62 NQIKILGNQGSF-LTGPSKLNDRADS-----RSEEVQLVFGLTANSDFHLLQGSLL 114
QY 140 TLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLONQKVEFKIDIV 199
DB 115 TLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLONQKVEFKIDIV 174
QY 200 V 200
DB 175 V 175

RESULT 65
US-08-630-172-1
Sequence 1, Application US/08630172
Patent No. 6060054

GENERAL INFORMATION:
APPLICANT: Staerz, Uwe
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
TITLE OF INVENTION: LYMPHOCYTE VETO
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, 35th floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,172
FILING DATE:
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-36
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-630-172-1

Query Match 54.4%; Score 716; DB 3; Length 138;
Best Local Similarity 100.0%; Pred. No. 2,2e-53;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 NKVLGKKGDVVELTCTASQKSIQFHWKNSNQIKILGNQGSFLTQKPSKLNDRADSRRS 85
DB 1 NKVLGKKGDVVELTCTASQKSIQFHWKNSNQIKILGNQGSFLTQKPSKLNDRADSRRS 60
QY 86 LMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANS DTHLHOGQSILTLLES 145
DB 61 LMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANS DTHLHOGQSILTLLES 120
QY 146 PPGSSPSVQCRSPRGKNI 163
DB 121 PPGSSPSVQCRSPRGKNI 138

RESULT 66
US-09-375-419-1
Sequence 1, Application US/09375419
Patent No. 6264950
GENERAL INFORMATION:
APPLICANT: Staerz, Uwe
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
TITLE OF INVENTION: LYMPHOCYTE VETO
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, 35th Floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/375,419
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,172
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-36
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-375-419-1

Query Match 54.4%; Score 716; DB 3; Length 138;
Best Local Similarity 100.0%; Pred. No. 2,2e-53;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 NKVLGKKGDVVELTCTASQKSIQFHWKNSNQIKILGNQGSFLTQKPSKLNDRADSRRS 85
DB 1 NKVLGKKGDVVELTCTASQKSIQFHWKNSNQIKILGNQGSFLTQKPSKLNDRADSRRS 60
QY 86 LMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANS DTHLHOGQSILTLLES 145
DB 61 LMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANS DTHLHOGQSILTLLES 120
QY 146 PPGSSPSVQCRSPRGKNI 163
DB 121 PPGSSPSVQCRSPRGKNI 138

RESULT 67
US-08-558-269-10
Sequence 10, Application US/08558269
Patent No. 5961973
GENERAL INFORMATION:
APPLICANT: Crea, Roberto
TITLE OF INVENTION: PATHOGEN-TARGETED BIOCATALYSTS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/558,269
FILING DATE: 13-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,800
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CRI-001CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-558-269-10

Query Match 45.2%; Score 595.5; DB 2; Length 376;
Best Local Similarity 92.2%; Pred. No. 1,4e-42;
Matches 118; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 27 KVLVGGKGDVVELTCTASQKSIQFHWKNSNQIKILGNQGSFLTQKPSKLNDRADSRRL 86
DB 3 KVLVGGKGDVVELTCTASQKSIQFHWKNSNQIKILGNQGSFLTQKPSKLNDRADSRRL 62
QY 87 WDOGNFPLIIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANS DTHLHOGQSILTLLES 146
DB 63 WDOGNFPLIIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANS DTHLHOGQSILTLLES 122

QY 147 P---GSSP 151
DB 123 DAELGMSF 130

RESULT 68

US-09-410-882-10
; Sequence 10, Application US/09410882
; Patent No. 6287561
; GENERAL INFORMATION:
; APPLICANT: Crea, Roberto
; TITLE OF INVENTION: PATHOGEN-TARGETED BIOCATALYSTS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/410,882
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,269
; FILING DATE: 13-NOV-1995
; APPLICATION NUMBER: US 07/847,800
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CRI-001CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-410-882-10

Query Match 45.2%; Score 595.5; DB 3; Length 376;
Best Local Similarity 92.2%; Pred. No. 1.4e-42;
Matches 118; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 27 KVLGKKGDTVELTCTASQKSIQFHWKNSNOIKILGNQSFITKGPSKLNDRADSRSL 86
DB 3 KVLGKKGDTVELTCTASQKSIQFHWKNSNOIKILGNQSFITKGPSKLNDRADSRSL 62
QY 87 WDQGNFPLIIKNLKIETSDTYICEVEQKEVQLVFGLTANSPTHLHOGQSLTLTLESP 146
DB 63 WDQGNFPLIIKNLKIETSDTYICEVEQKEVQLVFGLTANSPTHLHOGQSLTLTLESP 122
QY 147 P---GSSP 151
DB 123 DAELGMSF 130

RESULT 69

US-08-558-269-6
; Sequence 6, Application US/08558269
; Patent No. 5961973
; GENERAL INFORMATION:
; APPLICANT: Crea, Roberto
; TITLE OF INVENTION: PATHOGEN-TARGETED BIOCATALYSTS

NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/558,269
FILING DATE: 13-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,800
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CRI-001CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 383 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-558-269-6

Query Match 44.8%; Score 589.5; DB 2; Length 383;
Best Local Similarity 72.9%; Pred. No. 4.5e-42;
Matches 121; Conservative 9; Mismatches 23; Indels 13; Gaps 2;

QY 27 KVLGKKGDTVELTCTASQKSIQFHWKNSNOIKILGNQSFITKGPSKLNDRADSRSL 86
DB 3 KVLGKKGDTVELTCTASQKSIQFHWKNSNOIKILGNQSFITKGPSKLNDRADSRSL 62
QY 87 WDQGNFPLIIKNLKIETSDTYICEVEQKEVQLVFGLTANSPTHLHOGQSLTLTLESP 146
DB 63 WDQGNFPLIIKNLKIETSDTYICEVEQKEVQLVFGLTANSPTHLHOGQSLTLTLESP 117
QY 147 PGSSPSVOCSPRGKNIQGGKTLVSQLEIADSGTWTCTVLONQKR 192
DB 118 --STCGLRQYSQGPRIKGGLPADIA-----SHPWQAAIIFAGRR 155

RESULT 70

US-09-410-882-6
; Sequence 6, Application US/09410882
; Patent No. 6287561
; GENERAL INFORMATION:
; APPLICANT: Crea, Roberto
; TITLE OF INVENTION: PATHOGEN-TARGETED BIOCATALYSTS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/410,882
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/558,269
FILING DATE: 13-NOV-1995
APPLICATION NUMBER: US 07/847,800
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CRI-001CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 383 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-410-882-6

Query Match 44.8%; Score 589.5; DB 3; Length 383;
Best Local Similarity 72.9%; Pred. No. 4.5e-42;
Matches 121; Conservative 9; Mismatches 23; Indels 13; Gaps 2;

QY 27 KVVIGKKGVDELCTCTASQKKSIOFHWKNSNQIKILNQGSFLTKGSKLNDRADSRSL 86
DB 3 KVVIGKKGVDELCTCTASQKKSIOFHWKNSNQIKILNQGSFLTKGSKLNDRADSRSL 62

QY 87 WDOGPFLLIKNLKIEDSDTYICEVDQKEVQLVGLTANSTHLLQGSLLTLTLESP 146
DB 63 WDOGPFLLIKNLKIEDSDTYICEVDQKEVQLVGLTANSTHLLQGSLLTLTLESP 117

QY 147 PGSSPSVQCRSPRGKNIQCGKTLVLSQLELDQSGTWTCTVLANOKK 192
DB 118 --STCGLRQYSQPOFRKGGFLADIA-----SHPMQALPAKRR 155

RESULT 71
US-08-466-368-5
Sequence 5, Application US/08466368
Patent No. 6093539
GENERAL INFORMATION:
APPLICANT: Maddon, Paul J.
APPLICANT: Littman, Dan R.
APPLICANT: Chess, Leonard
APPLICANT: Axel, Richard
APPLICANT: Weles, Robin
APPLICANT: McDougal, J. S.
TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN
TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,368
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 24577-BI-B/JPW/AKC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: YES
FEATURE:
NAME/KEY: Active-site
LOCATION: 1..94
US-08-466-368-5

Query Match 37.5%; Score 494; DB 3; Length 94;
Best Local Similarity 100.0%; Pred. No. 8.5e-35;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 QGNKVVIGKKGVDELCTCTASQKKSIOFHWKNSNQIKILNQGSFLTKGSKLNDRADSR 83
DB 1 QGNKVVIGKKGVDELCTCTASQKKSIOFHWKNSNQIKILNQGSFLTKGSKLNDRADSR 60

QY 84 RSLMDQGNPFLLIKNLKIEDSDTYICEVDQKE 117
DB 61 RSLMDQGNPFLLIKNLKIEDSDTYICEVDQKE 94

RESULT 72
US-08-470-998-2
Sequence 2, Application US/08470998
Patent No. 6570000
GENERAL INFORMATION:
APPLICANT: Maddon, Paul J.
APPLICANT: Littman, Dan R.
APPLICANT: Chess, Leonard
APPLICANT: Axel, Richard
APPLICANT: Weles, Robin
APPLICANT: McDougal, J. S.
TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN
TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,998
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 24577-BI-B/JPW/AKC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: YES
FEATURE:
NAME/KEY: Active-site
LOCATION: 1..94
US-08-470-998-2

Query Match 37.5%; Score 494; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 8.5e-35;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

24 QGNKVVVGKGGKDPVELTCTASOKSIOFHMKNNOIKILNQGSLFKGSPKLNDRADSR 83
1 QGNKVVVGKGGKDPVELTCTASOKSIOFHMKNNOIKILNQGSLFKGSPKLNDRADSR 60

84 RSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEE 117
61 RSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEE 94

RESULT 73

US-08-328-500-10
Sequence 10; Application US/08328500
Patent No. 6673896
GENERAL INFORMATION:
APPLICANT: Maddon, Paul J.
APPLICANT: Axel, Richard
APPLICANT: Sweet, Richard W.
APPLICANT: Arthos, James
TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/328,500
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/24577-CY
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-328-500-10

Query Match 37.5%; Score 494; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 8.5e-35;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

24 QGNKVVVGKGGKDPVELTCTASOKSIOFHMKNNOIKILNQGSLFKGSPKLNDRADSR 83
1 QGNKVVVGKGGKDPVELTCTASOKSIOFHMKNNOIKILNQGSLFKGSPKLNDRADSR 60

84 RSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEE 117

61 RSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEE 94

RESULT 74
US-08-416-478A-8
Sequence 8; Application US/08416478A
Patent No. 5773578

GENERAL INFORMATION:
APPLICANT: Hercend, Thierry
APPLICANT: Tielebel, Frederic
TITLE OF INVENTION: New Proteins Produced By Human
Lymphocytes, DNA Sequences Encoding These Proteins And
Their Pharmaceutical And Biological Uses
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,478A
FILING DATE: 04-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,644
FILING DATE: 08-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: HERCEND=1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 457 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-416-478A-8

Query Match 37.2%; Score 490.5; DB 1; Length 457;
Best Local Similarity 48.8%; Pred. No. 1.4e-33;
Matches 102; Conservative 35; Mismatches 67; Indels 5; Gaps 3;

1 MNRGVFPRHL--LLVQLALPLPATQGNKVVVGKGGKDPVELTCTASOKSIOFHMKNNSG 58
1 MCRGFSFRHLPLLLQLSKLVLVYQKTVVLKREGSABLPESSTRSASASAMKSSDD 60

59 IKILNQGSLFKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEY 118
61 KTLIGYKNKLLIKGSLVSRFDSRKNAMERGSFPLINKLRMEDSQTVCLENNKEEV 120

119 QLLVFGULTANSDFHLQSGSLTITLES-PPGSSPSVQCRSPRKNIOGKTLVSQLEIQ 177
121 ELWVFRVTNPGFRLIQGSLTILIDSNSPKVSDPPIECRKSNIYKDSKAFSTHSLRIQ 180

178 DSGTWCTVLONQKVEF--KIDIVPRAS 204
181 DSGIHWCTVTLNOKKHSFMDKLSVLGFS 209

RESULT 75
US-08-474-988B-8

```

; Sequence 8, Application US/08474988B
; Patent No. 5874250
; GENERAL INFORMATION:
; APPLICANT: Hercend, Thierry
; APPLICANT: Triebel, Frederic
; TITLE OF INVENTION: NEW PROTEINS PRODUCED BY HUMAN
; TITLE OF INVENTION: LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
; TITLE OF INVENTION: THEIR PHARMACEUTICAL AND BIOLOGICAL USES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street, N.W., Suite 400
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,988B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 08/416,478
; FILING DATE: 04-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/854,644
; FILING DATE: 08-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: YUN, Allen C.
; REGISTRATION NUMBER: 37,971
; REFERENCE/DOCKET NUMBER: HERCEND=1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 457 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-474-988B-8

Query Match      37.2%; Score 490.5; DB 2; Length 457;
Best Local Similarity 48.8%; Pred. No. 1.4e-33;
Matches 102; Conservative 35; Mismatches 67; Indels 5; Gaps 3;

QY 1 MNRGVPRRH-LLVQLALLPAATQGNKVVLGKKDVTVELTCTASQKSIQFHKNSNQ 58
Db 1 MCRGFSFRHLLPLLQLSKLVLTQKTVLVGKGSALPCGSTRSRASFAWKSSDQ 60
QY 59 IKILNGSFLTKGPSKLNDRADRSRLMDQGNFPLIIKUKLIEDSDTYICEVDQKEEV 118
Db 61 KTLGYNKLLIKSLSLYSRFSRKNAMERGSFPLIINKLRMEDSQTYVCELENKKEEV 120
QY 119 QLLVFGLTANSDFHLLQGOSLTLTLES-PPGSSPSVOCRSPRKNIGQKTLVSQLELQ 177
Db 121 ELWVFRFTFNPGRLLQGOSLTLTLDSPKVSDBPIECKKSSNIVKDSAFTHSLRIQ 180
QY 178 DSGTWCTVLQONQKVEF--KIDIVPRAS 204
Db 181 DSGIWNCTVTLNOKGHSFDMKLSVLGFAS 209

RESULT 76
US-08-394-442B-8
; Sequence 8, Application US/08394442B
; Patent No. 5976877

```

```

; GENERAL INFORMATION:
; APPLICANT: Hercend, Thierry
; APPLICANT: Triebel, Frederic
; TITLE OF INVENTION: New Proteins Produced By Human
; TITLE OF INVENTION: Lymphocytes, DNA Sequences Encoding These Proteins And
; TITLE OF INVENTION: Their Pharmaceutical And Biological Uses
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street, N.W., Suite 400
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/394,442B
; FILING DATE: 24-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: HERCEND=2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 457 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-394-442B-8

Query Match      37.2%; Score 490.5; DB 2; Length 457;
Best Local Similarity 48.8%; Pred. No. 1.4e-33;
Matches 102; Conservative 35; Mismatches 67; Indels 5; Gaps 3;

QY 1 MNRGVPRRH-LLVQLALLPAATQGNKVVLGKKDVTVELTCTASQKSIQFHKNSNQ 58
Db 1 MCRGFSFRHLLPLLQLSKLVLTQKTVLVGKGSALPCGSTRSRASFAWKSSDQ 60
QY 59 IKILNGSFLTKGPSKLNDRADRSRLMDQGNFPLIIKUKLIEDSDTYICEVDQKEEV 118
Db 61 KTLGYNKLLIKSLSLYSRFSRKNAMERGSFPLIINKLRMEDSQTYVCELENKKEEV 120
QY 119 QLLVFGLTANSDFHLLQGOSLTLTLES-PPGSSPSVOCRSPRKNIGQKTLVSQLELQ 177
Db 121 ELWVFRFTFNPGRLLQGOSLTLTLDSPKVSDBPIECKKSSNIVKDSAFTHSLRIQ 180
QY 178 DSGTWCTVLQONQKVEF--KIDIVPRAS 204
Db 181 DSGIWNCTVTLNOKGHSFDMKLSVLGFAS 209

RESULT 77
US-08-332-562A-84
; Sequence 84, Application US/08332562A
; Patent No. 5985599
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, Ian F. C.
; APPLICANT: HOGARTH, Mark P. C.
; APPLICANT: HIBBS, Margaret L.
; APPLICANT: SCOTT, Bernadette M.
; APPLICANT: BONADONNA, Lisa
; APPLICANT: HULETT, Mark D.
; TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 136

```

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: 3000 K Street, N.W., Suite 500
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20007-5109
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/332,562A
;; FILING DATE: 31-OCT-1994
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/896,457
;; FILING DATE: 27-MAY-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BENT, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 54270/119/GRHA
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)672-5300
;; TELEFAX: (202)672-5399
;; TELEX: 904136
;; INFORMATION FOR SEQ ID NO: 84:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 80 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-332-562A-84

Query Match 30.2%; Score 398; DB 2; Length 80;
Best Local Similarity 98.7%; Pred. No. 9.3e-27;
Matches 77; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 123 FGLTANSDTHLQGSITLTLESPPGSSPSVOCSPRGKNIQGGKITLSVSLQELQDSGTW 182
DB 1 FGLTANSDTHLQGSITLTLESPPGSSPSVOCSPRGKNIQGGKITLSVSLQELQDSGTW 60

QY 183 TCTVLQNKQKVEFKIDIV 200
DB 61 TCTVLQNKQKVEFKIDIV 78

RESULT 78
US-09-050-861B-8
; Sequence 8, Application US/09050861B
; Patent No. 655314
; GENERAL INFORMATION:
; APPLICANT: Payan, Donald
; TITLE OF INVENTION: TOSO AS A TARGET FOR DRUG SCREENING
; FILE REFERENCE: RIGL-002CON
; CURRENT APPLICATION NUMBER: US/09/050,861B
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: US/09/651,150B
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: US 09/050,861
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 8
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-050-861B-8

Query Match 29.8%; Score 392; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 2.7e-26;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 VELTCTASQKSIQFHWKNSNQIKLGNQSFITKGPSKLANDRARRSLMDQGNPFLII 96
DB 1 VELTCTASQKSIQFHWKNSNQIKLGNQSFITKGPSKLANDRARRSLMDQGNPFLII 60

QY 97 KNLKIEDSDTYICE 110
DB 61 KNLKIEDSDTYICE 74

RESULT 79
US-09-100-409A-20
; Sequence 20, Application US/09100409A
; Patent No. 6030388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang YI
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; PREVENTION AND TREATMENT OF HIV INFECTION AND
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 amino acids
; TYPE: amino acid
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-409A-20

Query Match 23.8%; Score 314; DB 3; Length 61;
Best Local Similarity 100.0%; Pred. No. 8.5e-20;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 SKINDRADRRSLMDQGNPFLIIKNLKIEDSDTYICEVDOKEEYQVLVFGLTANSDDTHL 133
DB 1 SKINDRADRRSLMDQGNPFLIIKNLKIEDSDTYICEVDOKEEYQVLVFGLTANSDDTHL 60

QY 134 L 134
DB 61 L 61

RESULT 80
US-08-284-391B-35
; Sequence 35, Application US/08284391B
; Patent No. 5851828
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles

```
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-35

Query Match      21.9%; Score 288; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.3e-17;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      201 PRASALPAPPTGSAALPDPPQTASALPDPPASALPALALAVISFLIGLGIVACVLAARR 258
Db      1 PRASALPAPPTGSAALPDPPQTASALPDPPASALPALALAVISFLIGLGIVACVLAARR 58

RESULT 81
US-09-218-950-35
Sequence 35, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
```

```
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-35

Query Match      21.9%; Score 288; DB 3; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.3e-17;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      201 PRASALPAPPTGSAALPDPPQTASALPDPPASALPALALAVISFLIGLGIVACVLAARR 258
Db      1 PRASALPAPPTGSAALPDPPQTASALPDPPASALPALALAVISFLIGLGIVACVLAARR 58

RESULT 82
US-09-100-409A-10
Sequence 10, Application US/09100409A
Patent No. 6090388
GENERAL INFORMATION:
APPLICANT: Wang, Chang YI
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
```

```
REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-10

Query Match      17.3% Score 228; DB 3; Length 46;
Best Local Similarity 95.6%; Pred. No. 1.1e-12;
Matches 43; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 52 HMKNSNQIKILGNQSGFLTKGPSKLNDRADSRSLMDQGNCPILI 96
DB 2 HMKNSNQIKILGNQSGFLTKGPSKLNDRADSRSLMDQGNCPILI 46

RESULT 83
US-09-100-409A-34
Sequence 34, Application US/09100409A
Patent No. 6090388
GENERAL INFORMATION:
APPLICANT: Wang, Chang YI
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
PREVENTION AND TREATMENT OF HIV INFECTION AND
TITLE OF INVENTION: IMMUNE DISORDERS
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 81 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-34

Query Match      17.3% Score 228; DB 3; Length 81;
Best Local Similarity 95.6%; Pred. No. 2.4e-12;
Matches 43; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 52 HMKNSNQIKILGNQSGFLTKGPSKLNDRADSRSLMDQGNCPILI 96
DB 37 HMKNSNQIKILGNQSGFLTKGPSKLNDRADSRSLMDQGNCPILI 81

RESULT 84
```

```
US-08-630-172-7
Sequence 7, Application US/08630172
Patent No. 6060054
GENERAL INFORMATION:
APPLICANT: Staez, Uwe
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
TITLE OF INVENTION: LYMPHOCYTE VETO
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, 35th Floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,172
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-36
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 154 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-630-172-7

Query Match      16.1% Score 212; DB 3; Length 154;
Best Local Similarity 36.2%; Pred. No. 1.3e-10;
Matches 54; Conservative 18; Mismatches 19; Indels 58; Gaps 5;

QY 144 ESP-----GGSSPSVQCRS-----PRGNI-----163
DB 6 QSPHCTTVPGASVNTICSTSGGLRGIVLPOLGQPDIIYEDGVPTTDRFRGRIDS 65
QY 164 --GGKTLISVQLDLSGWTG-----TVLONOKVBEKI-DIVPRASA 205
DB 66 GSDNLTITMRLQLSDTGTTCQALTEVNVYGSGLVLTBERQSGMWRCSAPPASRA 125
QY 206 LPAPPTGSALPDPTASALPDPPASALP 234
DB 126 LPAPPTGSALPDPTASALPDPPASALP 154

RESULT 85
US-09-375-419-7
Sequence 7, Application US/09375419
Patent No. 6264950
GENERAL INFORMATION:
APPLICANT: Staez, Uwe
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
TITLE OF INVENTION: LYMPHOCYTE VETO
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, 35th Floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203
COMPUTER READABLE FORM:
```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

RESULT 86
 US-09-100-409A-2
 Sequence 2, Application US/09100409A
 Patent No. 6090388
 GENERAL INFORMATION:
 APPLICANT: Wang, Chang Yi
 TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
 PREVENTION AND TREATMENT OF HIV INFECTION AND
 TITLE OF INVENTION: IMMUNE DISORDERS
 NUMBER OF SEQUENCES: 64
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNEGAN
 STREET: 345 Park Avenue
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10154-0054
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version
 SOFTWARE: #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/100.409A
 FILING DATE:
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME:
 REGISTRATION NUMBER:

```

REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-2

Query Match          16.0%; Score 211; DB 3; Length 40;
Best Local Similarity 97.5%; Pred. No. 2,7e-11;
Matches 39; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db      1 HMKNSNQIKILGNQGSFLTGKPSKLTNDRADSRSLMDQGN 91
|||||
52 HMKNSNQIKILGNQGSFLTGKPSKLTNDRADSRSLMDQGN 91
|||||

RESULT 87
US-09-100-409A-4
Sequence 4, Application US/09100409A
Patent No. 6090388
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
PREVENTION AND TREATMENT OF HIV INFECTION AND
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-4

Query Match          16.0%; Score 211; DB 3; Length 42;
Best Local Similarity 97.5%; Pred. No. 2,7e-11;
Matches 39; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy      52 HMKNSNQIKILGNQGSFLTGKPSKLTNDRADSRSLMDQGN 91
|||||
Db      2 HMKNSNQIKILGNQGSFLTGKPSKLTNDRADSRSLMDQGN 91
|||||

```

US-09-100-409A-32
; Sequence 32, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-409A-32

Query Match 16.0%; Score 211; DB 3; Length 77;
Best Local Similarity 97.5%; Pred. No. 6.1e-11;
Matches 39; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 HMKNSQIKILNGSGSLTKGPSKLNDRADSRSLMDQGN 91
Db 37 HMKMNQIKILNGSGSLTKGPSKLNDRADSRSLMDQGN 76

RESULT 89
5510256-5
; Patent No. 5510256
; APPLICANT: KIRSCHNER, RICHARD J.;MOTT, JOHN E.;ECKENRODE,
; FRANCIS M.;BRUNNER, DAVID P.
; TITLE OF INVENTION: ELIMINATING INTERNAL INITIATION OF
; SOLUBLE CD4 GENE
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,995
; FILING DATE: 29-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 13,828
; FILING DATE: 02-FEB-1993
; APPLICATION NUMBER: 562,861
; FILING DATE: 06-AUG-1990
; SEQ ID NO:5:
; LENGTH: 41
; 5510256-5

Query Match 15.2%; Score 200; DB 6; Length 41;
Best Local Similarity 97.6%; Pred. No. 2.3e-10;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 95 IIKNLIKEDSDPYICEVEDEKKEVOLLVGLTANSDTHLQ 135
Db 1 IIKNKKIEDSDPYICEVEDEKKEVOLLVGLTANSDTHLQ 41

RESULT 90
US-08-332-562A-85
; Sequence 85, Application US/08332562A
; Patent No. 5985599
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, Ian F.C.
; APPLICANT: HOGARTH, Mark P.
; APPLICANT: HIBBS, Margaret L.
; APPLICANT: SCOTT, Bernadette M.
; APPLICANT: BONADONNA, Lisa
; APPLICANT: HULETT, Mark D.
; TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 136
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,562A
; FILING DATE: 31-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/896,457
; FILING DATE: 27-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 54270/119/GRHA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 81 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-332-562A-85

Query Match 12.8%; Score 168.5; DB 2; Length 81;
Best Local Similarity 50.7%; Pred. No. 2.6e-07;
Matches 36; Conservative 11; Mismatches 23; Indels 1; Gaps 1;

QY 123 FGLTANSDTHLQGGSLTTLTLES-PPGSSPVQCRSPRKNIQGGTLLSVSQLEQDSGT 181
Db 1 FKVTFSPTGSLTQGGSLTTLTLDNSKVSNPTECKHKKGKVSQSVLSMSNLRVQDSPF 60

QY 182 WTCVTVLQDNKK 192
Db 61 WNCVTVLQDNKK 71

RESULT 91
US-08-332-562A-86
; Sequence 86, Application US/08332562A
; Patent No. 5985599
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, Ian F.C.

APPLICANT: HOGARTH, Mark P.
APPLICANT: HIBBS, Margaret L.
APPLICANT: SCOTT, Bernadette M.
APPLICANT: BONADONNA, Lisa
APPLICANT: HULETT, Mark D.
TITLE OF INVENTION: PC RECEPTOR FOR IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,562A
FILING DATE: 31-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,457
FILING DATE: 27-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 54270/119/GRHA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 81 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-332-562A-86

Query Match 12.7%; Score 167.5; DB 2; Length 81;
Best Local Similarity 44.9%; Pred. No. 3.2e-07;
Matches 35; Conservative 10; Mismatches 32; Indels 1; Gaps 1;

QY 123 FGLTANSDTHLLQGSILTLTLES-PPGSSPSVQCRSPRGKNIQCGKTLVSQLELDPSGT 181
DB 1 FRYTFNPGTRLLQGSILTLTLDSPKVSDDPIECKHSSNIVKSKAFSTHSLRIDPSGI 60

QY 182 WTCVTLQOKKVEFKIDI 199
DB 61 WNCVTTLNOKKHSFDMKL 78

RESULT 92
US-09-100-409A-11
Sequence 11, Application US/09100409A
Patent No. 6090388
GENERAL INFORMATION:
APPLICANT: Wang, Chang YI
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
PREVENTION AND TREATMENT OF HIV INFECTION AND
TITLE OF INVENTION: IMMUNE DISORDERS
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-11

Query Match 12.7%; Score 167; DB 3; Length 34;
Best Local Similarity 97.0%; Pred. No. 1.1e-07;
Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 64 NQGSFLTGPSTKLNDRADSRSLMDQGNFPLII 96
DB 2 NQGSFLTGPSTKLNDRADSRSLMDQGNCFLLI 34

RESULT 93
US-09-100-409A-33
Sequence 33, Application US/09100409A
Patent No. 6090388
GENERAL INFORMATION:
APPLICANT: Wang, Chang YI
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
PREVENTION AND TREATMENT OF HIV INFECTION AND
TITLE OF INVENTION: IMMUNE DISORDERS
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 69 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-09-100-409A-33

Query Match 12.7%; Score 167; DB 3; Length 69;
Best Local Similarity 97.0%; Pred. No. 2.8e-07;
Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 64 NQGSFLTKGPSKLNDRADSRSLMDQGNFLLII 96
Db 37 NQGSFLTKGPSKLNDRADSRSLMDQGNFLLII 69

RESULT 94

US-09-100-409A-19
Sequence 19, Application US/09100409A
Patent No. 6090388
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
PREVENTION AND TREATMENT OF HIV INFECTION AND
TITLE OF INVENTION: IMMUNE DISORDERS
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-19

Query Match 11.8%; Score 156; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 7.6e-07;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 NQGSFLTKGPSKLNDRADSRSLMDQGNF 92
Db 1 NQGSFLTKGPSKLNDRADSRSLMDQGNF 29

RESULT 95
US-09-100-414B-92
Sequence 92, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.

STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-92

Query Match 11.5%; Score 151.5; DB 3; Length 47;
Best Local Similarity 66.0%; Pred. No. 3.5e-06;
Matches 33; Conservative 5; Mismatches 5; Indels 7; Gaps 2;

OY 44 SQKSIQIFHWKNSNQIKILG--NQGSFLTKGPSKLNDRADSRSLMDQGN 91
Db 2 SEIKGVYH-----KIEGIGGCGNQGSFLTKGPSKLNDRADSRSLMDQGN 46

RESULT 96
US-09-303-323-92
Sequence 92, Application US/09303323
Patent No. 6228987
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849

Search completed: August 3, 2004, 13:17:01
 Job time : 9.59249 secs

NAME: Maria H. Lin
 REGISTRATION NUMBER: 29,323
 REFERENCE/DOCKET NUMBER: 1151-4157
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-758-4800
 TELEFAX: 212-751-6849
 INFORMATION FOR SEQ ID NO: 91:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 30 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-09-100-414B-91

Query Match 11.4%; Score 150; DB 3; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 NQGSFLTKGPSKLNDRADSRSLWDQGN 91
 DB 2 NQGSFLTKGPSKLNDRADSRSLWDQGN 29

RESULT 100

US-09-100-409A-5
 Sequence 5, Application US/09100409A
 Patent No. 6090388

GENERAL INFORMATION:

APPLICANT: Wang, Chang Yi

TITLE OF INVENTION: PEPTIDE COMPOSITION FOR

TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND

TITLE OF INVENTION: IMMUNE DISORDERS

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0054

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/100,409A

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME:

REGISTRATION NUMBER:

REFERENCE/DOCKET NUMBER: 1151-4154

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-758-4800

TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-100-409A-5

Query Match 11.4%; Score 150; DB 3; Length 30;

Best Local Similarity 100.0%; Pred. No. 2.6e-06;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 NQGSFLTKGPSKLNDRADSRSLWDQGN 91
 DB 2 NQGSFLTKGPSKLNDRADSRSLWDQGN 29

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 13:01:34 ; Search time 13.331 Seconds
(without alignments) 1754.300 Million cell updates/sec

Title: SEQ5
Perfect score: 2335
Sequence: 1 MNRGVPRHLLVLQLALP.....VISFLIGLGIVACVLRTR 453

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 125 summaries

Database :

1: Issued Patents AA: *
2: /cgn2_6/ptodata/2/1aa/5A_COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------------|
| 1 | 2046 | 88.0 | 458 | 3 | US-08-466-368-4 |
| 2 | 2043.5 | 87.9 | 630 | 4 | US-08-472-888A-6 |
| 3 | 2040 | 87.7 | 462 | 2 | US-08-417-495-5 |
| 4 | 2040 | 87.7 | 462 | 2 | US-08-284-391B-5 |
| 5 | 2040 | 87.7 | 462 | 3 | US-09-218-950-5 |
| 6 | 2040 | 87.7 | 462 | 5 | PCT-US92-01785-5 |
| 7 | 2040 | 87.7 | 462 | 5 | PCT-US95-00454-5 |
| 8 | 2038.5 | 87.7 | 575 | 2 | US-08-417-495-4 |
| 9 | 2038.5 | 87.7 | 575 | 2 | US-08-284-391B-4 |
| 10 | 2038.5 | 87.7 | 575 | 2 | US-09-218-950-4 |
| 11 | 2038.5 | 87.7 | 575 | 5 | PCT-US92-01785-4 |
| 12 | 2038.5 | 87.7 | 575 | 5 | PCT-US95-00454-4 |
| 13 | 2038 | 87.7 | 457 | 4 | US-08-328-500-9 |
| 14 | 2037 | 87.6 | 398 | 3 | US-08-284-391B-29 |
| 15 | 2037 | 87.6 | 398 | 3 | US-09-218-950-29 |
| 16 | 2037 | 87.6 | 532 | 2 | US-08-417-495-6 |
| 17 | 2037 | 87.6 | 532 | 2 | US-08-284-391B-6 |
| 18 | 2037 | 87.6 | 532 | 3 | US-09-218-950-6 |
| 19 | 2037 | 87.6 | 532 | 5 | PCT-US92-01785-6 |
| 20 | 2037 | 87.6 | 532 | 5 | PCT-US95-00454-6 |
| 21 | 2032 | 87.4 | 458 | 3 | US-09-039-555B-15 |
| 22 | 2024 | 87.1 | 458 | 4 | US-09-517-605-3 |
| 23 | 2017 | 86.8 | 402 | 1 | US-08-236-311-1 |
| 24 | 2017 | 86.8 | 402 | 3 | US-08-457-918-1 |
| 25 | 2006 | 86.3 | 458 | 6 | 5223394-7 |
| 26 | 2001 | 86.1 | 394 | 3 | US-08-466-368-2 |
| 27 | 2001 | 86.1 | 394 | 4 | US-08-328-500-2 |

| | | | | | | |
|-----|--------|------|-----|---|-------------------|--------------------|
| 28 | 1951 | 83.9 | 394 | 6 | 5223418-2 | Patent No. 5223418 |
| 29 | 1909 | 82.1 | 433 | 2 | US-08-867-149-1 | Sequence 1, Appli |
| 30 | 1904 | 81.9 | 433 | 2 | US-08-808-374-1 | Sequence 1, Appli |
| 31 | 1904 | 81.9 | 433 | 3 | US-09-100-409A-1 | Sequence 1, Appli |
| 32 | 1904 | 81.9 | 434 | 1 | US-08-236-311-4 | Sequence 4, Appli |
| 33 | 1904 | 81.9 | 434 | 3 | US-08-457-918-4 | Sequence 4, Appli |
| 34 | 1712 | 73.6 | 433 | 6 | 5171838-13 | Patent No. 5171838 |
| 35 | 1368 | 58.8 | 318 | 6 | 5223394-11 | Patent No. 5223394 |
| 36 | 1363 | 58.6 | 295 | 6 | 5223394-9 | Patent No. 5223394 |
| 37 | 1050.5 | 45.2 | 310 | 3 | US-08-477-460B-6 | Sequence 6, Appli |
| 38 | 1050.5 | 45.2 | 310 | 3 | US-08-379-516-6 | Sequence 6, Appli |
| 39 | 1050.5 | 45.2 | 310 | 3 | US-09-329-916-6 | Sequence 6, Appli |
| 40 | 1050.5 | 45.2 | 310 | 3 | US-08-485-372A-6 | Sequence 6, Appli |
| 41 | 1050.5 | 45.2 | 310 | 4 | US-09-409-006A-6 | Sequence 6, Appli |
| 42 | 1050.5 | 45.2 | 310 | 4 | US-08-484-681-6 | Sequence 6, Appli |
| 43 | 1050.5 | 45.2 | 310 | 5 | PCT-US93-07422-6 | Sequence 6, Appli |
| 44 | 1046 | 45.0 | 530 | 3 | US-08-477-460B-4 | Sequence 4, Appli |
| 45 | 1046 | 45.0 | 530 | 3 | US-08-379-516-4 | Sequence 4, Appli |
| 46 | 1046 | 45.0 | 530 | 3 | US-09-329-916-4 | Sequence 4, Appli |
| 47 | 1046 | 45.0 | 530 | 3 | US-08-485-372A-4 | Sequence 4, Appli |
| 48 | 1046 | 45.0 | 530 | 4 | US-09-409-006A-4 | Sequence 4, Appli |
| 49 | 1046 | 45.0 | 530 | 4 | US-08-484-681-4 | Sequence 4, Appli |
| 50 | 1046 | 45.0 | 530 | 5 | PCT-US93-07422-4 | Sequence 4, Appli |
| 51 | 1045 | 44.9 | 432 | 3 | US-08-477-460B-2 | Sequence 2, Appli |
| 52 | 1045 | 44.9 | 432 | 3 | US-08-379-516-2 | Sequence 2, Appli |
| 53 | 1045 | 44.9 | 432 | 3 | US-09-329-916-2 | Sequence 2, Appli |
| 54 | 1045 | 44.9 | 432 | 3 | US-08-485-372A-2 | Sequence 2, Appli |
| 55 | 1045 | 44.9 | 432 | 4 | US-09-409-006A-2 | Sequence 2, Appli |
| 56 | 1045 | 44.9 | 432 | 4 | US-08-484-681-2 | Sequence 2, Appli |
| 57 | 1045 | 44.9 | 432 | 5 | PCT-US93-07422-2 | Sequence 2, Appli |
| 58 | 1041 | 44.8 | 203 | 2 | US-08-284-391B-31 | Sequence 31, Appli |
| 59 | 1041 | 44.8 | 203 | 3 | US-09-218-950-31 | Sequence 31, Appli |
| 60 | 999 | 43.0 | 457 | 1 | US-08-416-478A-8 | Sequence 8, Appli |
| 61 | 999 | 43.0 | 457 | 2 | US-08-474-968B-8 | Sequence 8, Appli |
| 62 | 999 | 43.0 | 457 | 2 | US-08-394-442B-8 | Sequence 8, Appli |
| 63 | 912.5 | 39.2 | 410 | 3 | US-08-630-172-17 | Sequence 17, Appli |
| 64 | 912.5 | 39.2 | 410 | 3 | US-09-375-419-17 | Sequence 17, Appli |
| 65 | 911 | 39.2 | 903 | 1 | US-08-021-601-12 | Sequence 12, Appli |
| 66 | 911 | 39.2 | 903 | 1 | US-08-082-849B-12 | Sequence 12, Appli |
| 67 | 911 | 39.2 | 903 | 5 | PCT-US94-01624-12 | Sequence 12, Appli |
| 68 | 716 | 30.8 | 138 | 3 | US-08-630-172-1 | Sequence 1, Appli |
| 69 | 716 | 30.8 | 138 | 3 | US-09-375-419-1 | Sequence 1, Appli |
| 70 | 603 | 25.9 | 376 | 2 | US-08-558-269-10 | Sequence 10, Appli |
| 71 | 603 | 25.9 | 376 | 2 | US-09-410-882-10 | Sequence 10, Appli |
| 72 | 589.5 | 25.4 | 383 | 2 | US-08-558-269-6 | Sequence 6, Appli |
| 73 | 589.5 | 25.4 | 383 | 3 | US-09-410-882-6 | Sequence 6, Appli |
| 74 | 494 | 21.2 | 94 | 4 | US-08-466-368-5 | Sequence 5, Appli |
| 75 | 494 | 21.2 | 94 | 4 | US-08-470-998-2 | Sequence 2, Appli |
| 76 | 494 | 21.2 | 94 | 4 | US-08-328-500-10 | Sequence 10, Appli |
| 77 | 406 | 17.5 | 80 | 2 | US-08-332-562A-84 | Sequence 84, Appli |
| 78 | 392 | 16.9 | 74 | 4 | US-09-050-861B-8 | Sequence 8, Appli |
| 79 | 391 | 16.8 | 83 | 2 | US-08-332-562A-87 | Sequence 87, Appli |
| 80 | 314 | 13.5 | 61 | 3 | US-09-100-409A-20 | Sequence 20, Appli |
| 81 | 288 | 12.4 | 58 | 2 | US-08-284-391B-35 | Sequence 35, Appli |
| 82 | 288 | 12.4 | 58 | 3 | US-09-218-950-35 | Sequence 35, Appli |
| 83 | 228 | 9.8 | 46 | 3 | US-09-100-409A-10 | Sequence 10, Appli |
| 84 | 228 | 9.8 | 46 | 3 | US-09-100-409A-34 | Sequence 34, Appli |
| 85 | 211 | 9.1 | 42 | 3 | US-09-100-409A-2 | Sequence 2, Appli |
| 86 | 211 | 9.1 | 42 | 3 | US-09-100-409A-4 | Sequence 4, Appli |
| 87 | 211 | 9.1 | 77 | 3 | US-09-100-409A-32 | Sequence 32, Appli |
| 88 | 200 | 8.6 | 41 | 6 | 5510256-5 | Patent No. 5510256 |
| 89 | 184.5 | 7.9 | 154 | 3 | US-08-630-172-7 | Sequence 7, Appli |
| 90 | 184.5 | 7.9 | 154 | 3 | US-09-375-419-7 | Sequence 7, Appli |
| 91 | 174.5 | 7.5 | 81 | 2 | US-08-332-562A-85 | Sequence 85, Appli |
| 92 | 173.5 | 7.5 | 81 | 2 | US-08-332-562A-86 | Sequence 86, Appli |
| 93 | 167 | 7.2 | 34 | 3 | US-09-100-409A-11 | Sequence 11, Appli |
| 94 | 167 | 7.2 | 69 | 3 | US-09-100-409A-33 | Sequence 33, Appli |
| 95 | 156 | 6.7 | 29 | 3 | US-09-100-409A-13 | Sequence 13, Appli |
| 96 | 155 | 6.7 | 735 | 5 | PCT-US93-00031-13 | Sequence 13, Appli |
| 97 | 155 | 6.7 | 736 | 5 | PCT-US93-00031-15 | Sequence 15, Appli |
| 98 | 155 | 6.7 | 739 | 4 | US-08-482-073-6 | Sequence 6, Appli |
| 99 | 155 | 6.7 | 739 | 5 | PCT-US93-00031-9 | Sequence 9, Appli |
| 100 | 154 | 6.6 | 662 | 1 | US-08-261-304-7 | Sequence 7, Appli |

101 151.5 6.5 47 3 US-09-100-414B-92 Sequence 92, Appl
102 151.5 6.5 47 3 US-09-303-323-92 Sequence 92, Appl
103 151.5 6.5 47 4 US-09-770-014-92 Sequence 92, Appl
104 150.5 6.5 828 1 US-08-261-304-2 Sequence 2, Appl
105 150 6.5 28 3 US-09-100-409A-3 Sequence 3, Appl
106 150 6.5 30 3 US-09-100-414B-91 Sequence 91, Appl
107 150 6.5 30 3 US-09-100-409A-5 Sequence 5, Appl
108 150 6.5 30 3 US-09-303-323-91 Sequence 91, Appl
109 150 6.5 30 4 US-09-770-014-91 Sequence 91, Appl
110 150 6.5 47 3 US-09-100-414B-93 Sequence 93, Appl
111 150 6.5 47 3 US-09-100-414B-94 Sequence 94, Appl
112 150 6.5 47 3 US-09-100-409A-60 Sequence 60, Appl
113 150 6.5 47 3 US-09-303-323-93 Sequence 93, Appl
114 150 6.5 47 3 US-09-303-323-94 Sequence 94, Appl
115 150 6.5 47 4 US-09-770-014-93 Sequence 93, Appl
116 150 6.5 47 4 US-09-770-014-94 Sequence 94, Appl
117 150 6.5 50 3 US-09-100-409A-37 Sequence 37, Appl
118 144 6.2 64 3 US-09-100-409A-35 Sequence 35, Appl
119 143 6.2 647 4 US-09-009-490A-91 Sequence 91, Appl
120 143 6.2 647 4 US-08-482-073-5 Sequence 5, Appl
121 143 6.2 647 5 PCT-US93-00031-11 Sequence 11, Appl
122 139 6.0 26 5 PCT-US93-00031-23 Patent No. 5171838
123 138 5.9 647 5 PCT-US93-00031-23 Sequence 23, Appl
124 137.5 5.9 642 1 US-08-217-299-1 Sequence 1, Appl
125 137.5 5.9 698 2 US-08-602-725-36 Sequence 36, Appl

ALIGNMENTS

RESULT 1
US-08-466-368-4
Sequence 4, Application US/08466368

Patent No. 6093539

GENERAL INFORMATION:

APPLICANT: Madden, Paul J.

APPLICANT: Littman, Dan R.

APPLICANT: Chess, Leonard

APPLICANT: Axel, Richard

APPLICANT: Weis, Robin

APPLICANT: McDougal, J. S.

TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN

TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,368

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 24577-E1-B/JPW/AKC

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-278-0400

TELEFAX: 212-391-0525

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 458 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-466-368-4
Query Match 88.0%; Score 2046; DB 3; Length 458;
Best Local Similarity 89.6%; Pred. No. 6.2e-163;
Matches 407; Conservative 2; Mismatches 15; Indels 30; Gaps 2;

QY 1 MNRGVFRLHLLVQLALPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHKNSNOIK 60
DB 1 MNRGVFRLHLLVQLALPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHKNSNOIK 60
QY 61 ILNQGSLFTKGSKLNDRADSRRLMDQGFLLIKNLKIEDSDYTCVEDEQKEVQL 120
DB 61 ILNQGSLFTKGSKLNDRADSRRLMDQGFLLIKNLKIEDSDYTCVEDEQKEVQL 120
QY 121 LVGGLTANSTHLLQGSLLTLESPPGSSPSVQCSPRGKNIQGGKTLISVQLDQSG 180
DB 121 LVGGLTANSTHLLQGSLLTLESPPGSSPSVQCSPRGKNIQGGKTLISVQLDQSG 180
QY 181 TWCTVLOQKVEFKIDIVLAFQKASSIVYKKEGBOVEFPLAFTVEKLTGSGELMW 240
DB 181 TWCTVLOQKVEFKIDIVLAFQKASSIVYKKEGBOVEFPLAFTVEKLTGSGELMW 240
QY 241 QABRASSSKSWITFDLKNKEVSVKRYTQPKLQMGKKLPLHLTLPOLPYAGSGNITLA 300
DB 241 QABRASSSKSWITFDLKNKEVSVKRYTQPKLQMGKKLPLHLTLPOLPYAGSGNITLA 300
QY 301 LEAKTGKLEHGVNLYVMRATQLOKNTCEYMGPTSPKMLSTLENKEAKVSKREKPYWV 360
DB 301 LEAKTGKLEHGVNLYVMRATQLOKNTCEYMGPTSPKMLSTLENKEAKVSKREKPYWV 360
QY 361 LNEBAGMOCCLSDSQVLLSNIKYLPWSTPEVHPASALPAPPGSALPDQOTASALP 420
DB 361 LNEBAGMOCCLSDSQVLLSNIKYLPWSTPEVHPASALPAPPGSALPDQOTASALP 420
QY 421 DPAASALPAAALVIFSLGLGLGV-ACVLARR 453
DB 397 -----MALIVLGGVAGLLFIIGIFFCVRCRR 425

RESULT 2

US-08-472-888A-6
Sequence 6, Application US/08472888A

Patent No. 6613746

GENERAL INFORMATION:

APPLICANT: Weis, Brian

APPLICANT: Seed, Gerald

TITLE OF INVENTION: ACP-ANTIBODY FUSION PROTEINS

TITLE OF INVENTION: AND RELATED MOLECULES AND METHODS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark & Elbing LLP

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,888A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/618,314

FILING DATE: 23-NOV-1990

ATTORNEY/AGENT INFORMATION:

NAME: Elbing, Karen L.

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/258001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200
 TELEFAX: 617-428-7045
 TELEEX:
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 630 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-472-888A-6

Query Match 87.9%; Score 2043.5; DB 4; Length 630;
 Best Local Similarity 95.0%; Pred. No. 1,66-162;
 Matches 401; Conservative 2; Mismatches 12; Indels 7; Gaps 1;

```

QY 1 MNRGVPRHLLVLVQLALPPATQGNKVLGKKGDVELTCTASQKSIQFHWKNSNOIK 60
DB 1 MNRGVPRHLLVLVQLALPPATQGNKVLGKKGDVELTCTASQKSIQFHWKNSNOIK 60
QY 61 ILNGGSLFTKGPSKLNDRADSRSLMDQGNPFLIIKNLKIEDSDTYICEVEDQKEEYQL 120
DB 61 ILNGGSLFTKGPSKLNDRADSRSLMDQGNPFLIIKNLKIEDSDTYICEVEDQKEEYQL 120
QY 121 LVFGLTANSDDTHLQGGSLTLTLESPPSSPSVOCSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDDTHLQGGSLTLTLESPPSSPSVOCSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCVLOQOKKVEFKIDIVLAFQKASSIYKKEGEQVEFSPLAFVTEKLTSGGELMW 240
DB 181 TWTCVLOQOKKVEFKIDIVLAFQKASSIYKKEGEQVEFSPLAFVTEKLTSGGELMW 240
QY 241 QAERASSSSSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300
DB 241 QAERASSSSSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300
QY 301 LEAKTGKLGHOEVNLYVMRATQLOKNTLCEVWGPTSPKMLSLKENKAKYSKREKPYWV 360
DB 301 LEAKTGKLGHOEVNLYVMRATQLOKNTLCEVWGPTSPKMLSLKENKAKYSKREKPYWV 360
QY 361 LNPEAGMQCLSDSGOVLLESNIKVLPTWSTPVHPRASALPAPPTGSALPDPOATASALP 420
DB 361 LNPEAGMQCLSDSGOVLLESNIKVLPTWSTPVHPRASALPAPPTGSALPDPOATASALP 420
QY 421 DP 422
DB 421 DP 422
QY 414 CP 415
DB 414 CP 415

```

RESULT 3
 US-08-417-495-5
 ; Sequence 5, Application US/08417495
 ; Patent No. 5843728
 ; GENERAL INFORMATION:
 ; APPLICANT: Seed, Brian et al.
 ; TITLE OF INVENTION: Redirection of Cellular Immunity by Chimeraab
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 50Z or 55SX
 ; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 ; SOFTWARE: Wordperfect (Version 5.0)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/417,495
 ; FILING DATE:

CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/203,866
 FILING DATE:
 APPLICATION NUMBER: US/07/847,566
 FILING DATE:
 APPLICATION NUMBER: 07/665,961
 FILING DATE: March 7, 1991
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 00786/119002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEEX: 200154
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 462 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-417-495-5

Query Match 87.7%; Score 2040; DB 2; Length 462;
 Best Local Similarity 91.4%; Pred. No. 2e-162;
 Matches 405; Conservative 1; Mismatches 9; Indels 28; Gaps 2;

```

QY 1 MNRGVPRHLLVLVQLALPPATQGNKVLGKKGDVELTCTASQKSIQFHWKNSNOIK 60
DB 1 MNRGVPRHLLVLVQLALPPATQGNKVLGKKGDVELTCTASQKSIQFHWKNSNOIK 60
QY 61 ILNGGSLFTKGPSKLNDRADSRSLMDQGNPFLIIKNLKIEDSDTYICEVEDQKEEYQL 120
DB 61 ILNGGSLFTKGPSKLNDRADSRSLMDQGNPFLIIKNLKIEDSDTYICEVEDQKEEYQL 120
QY 121 LVFGLTANSDDTHLQGGSLTLTLESPPSSPSVOCSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDDTHLQGGSLTLTLESPPSSPSVOCSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCVLOQOKKVEFKIDIVLAFQKASSIYKKEGEQVEFSPLAFVTEKLTSGGELMW 240
DB 181 TWTCVLOQOKKVEFKIDIVLAFQKASSIYKKEGEQVEFSPLAFVTEKLTSGGELMW 240
QY 241 QAERASSSSSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300
DB 241 QAERASSSSSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300
QY 301 LEAKTGKLGHOEVNLYVMRATQLOKNTLCEVWGPTSPKMLSLKENKAKYSKREKPYWV 360
DB 301 LEAKTGKLGHOEVNLYVMRATQLOKNTLCEVWGPTSPKMLSLKENKAKYSKREKPYWV 360
QY 361 LNPEAGMQCLSDSGOVLLESNIKVLPTWSTPVHPRASALPAPPTGSALPDPOATASALP 420
DB 361 LNPEAGMQCLSDSGOVLLESNIKVLPTWSTPVHPRASALPAPPTGSALPDPOATASALP 420
QY 421 DP 422
DB 421 DP 422
QY 400 -----LCYILDAIILFLYGIVL 415
DB 400 -----LCYILDAIILFLYGIVL 415

```

RESULT 4
 US-08-284-391B-5
 ; Sequence 5, Application US/08284391B
 ; Patent No. 5851828
 ; GENERAL INFORMATION:
 ; APPLICANT: Seed, Brian
 ; APPLICANT: Banapour, Babak
 ; APPLICANT: Romeo, Charles
 ; APPLICANT: Kolanus, Waldemar
 ; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
 ; CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS

NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-5

Query Match 87.7%; Score 2040; DB 2; Length 462;
Best Local Similarity 91.4%; Pred. No. 2e-162;
Matches 405; Conservative 1; Mismatches 9; Indels 28; Gaps 2;
QY 1 MNRGVPFRHLLVQLALPPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHWKNSNOIK 60
DB 1 MNRGVPFRHLLVQLALPPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHWKNSNOIK 60
QY 61 ILNGQSFLLTGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILNGQSFLLTGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGSLLTTLSPGSSPSVQCRSPRKNIOGGKTLVSQLELDGSG 180
DB 121 LVFGLTANSDTHLLQGSLLTTLSPGSSPSVQCRSPRKNIOGGKTLVSQLELDGSG 180
QY 181 TWCTCVLONOKKVEFKDIVLAFOKASIIYKKEGVEVSFLAFVETLTSGLLW 240
DB 181 TWCTCVLONOKKVEFKDIVLAFOKASIIYKKEGVEVSFLAFVETLTSGLLW 240
QY 241 QAERASSSSKSWITFDLKNKEVSVKRVTDPRLOMGKLLPLHLTPQALPYAGSGNLTIA 300
DB 241 QAERASSSSKSWITFDLKNKEVSVKRVTDPRLOMGKLLPLHLTPQALPYAGSGNLTIA 300
QY 301 LEATGKLGHEVNVLMVRATOLQKNTLCEVWGPTSPKMLSLKLENKAVSKKEKPVW 360
DB 301 LEATGKLGHEVNVLMVRATOLQKNTLCEVWGPTSPKMLSLKLENKAVSKKEKPVW 360
QY 361 LNPBAGMWQCLLSQGVLLSNNIKVLPWTSTVHPBASALPAPPTGSALPDPOTTAALP 420
DB 361 LNPBAGMWQCLLSQGVLLSNNIKVLPWTSTVHPBASALPAPPTGSALPDPOTTAALP 420
QY 361 LNPBAGMWQCLLSQGVLLSNNIKVLPWTSTVHPBASALPAPPTGSALPDPOTTAALP 399
DB 361 LNPBAGMWQCLLSQGVLLSNNIKVLPWTSTVHPBASALPAPPTGSALPDPOTTAALP 399

QY 421 DPPASALPAAALAVISFLGLGL 443
DB 400 -----LCYILDALIFLVGIVL 415

RESULT 5
US-09-218-950-5
Sequence 5, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-5

Query Match 87.7%; Score 2040; DB 3; Length 462;
Best Local Similarity 91.4%; Pred. No. 2e-162;
Matches 405; Conservative 1; Mismatches 9; Indels 28; Gaps 2;
QY 1 MNRGVPFRHLLVQLALPPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHWKNSNOIK 60
DB 1 MNRGVPFRHLLVQLALPPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHWKNSNOIK 60
QY 61 ILNGQSFLLTGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILNGQSFLLTGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGSLLTTLSPGSSPSVQCRSPRKNIOGGKTLVSQLELDGSG 180
DB 121 LVFGLTANSDTHLLQGSLLTTLSPGSSPSVQCRSPRKNIOGGKTLVSQLELDGSG 180

Db 121 LVFGLTANSDBTHLQGSGLTTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCVTLONOKKVEFKIDIVLAFQKASSIYVKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
Db 181 TWTCVTLONOKKVEFKIDIVLAFQKASSIYVKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSSKSWITFDLKNKEVSKRVTPDKLQMGKULPLHLTLPOALPOYAGSGLTLA 300
Db 241 QAERASSSSKSWITFDLKNKEVSKRVTPDKLQMGKULPLHLTLPOALPOYAGSGLTLA 300
QY 301 LEAKTGKLGHOEVNLYVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPYWV 360
Db 301 LEAKTGKLGHOEVNLYVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPYWV 360
QY 361 INPEAGMWOCILSDSGQVLESNIKVLPTWSTPVHPRASALPAPPTGSALPDPOTASALP 420
Db 361 INPEAGMWOCILSDSGQVLESNIKVLPTWSTPVH-----ADPQ----- 399
QY 421 DPPASALPALAVISFLGLGL 443
Db 400 -----LCYILDALFLYGLVL 415
RESULT 6
PCT-US92-01785-5
Sequence 5, Application PC/TUS9201785
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
TITLE OF INVENTION: Chimeras
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-01785-5
Query Match 87.7%; Score 2040; DB 5; Length 462;
Best Local Similarity 91.4%; Pred No. 2e-162;
Matches 405; Conservative 1; Mismatches 9; Indels 28; Gaps 2;
QY 1 NNRGVFPRHLVLVLQALPLPATGKNKVLGKKGDVTELTCTASQKSIQFMKNSNDIK 60
Db 1 NNRGVFPRHLVLVLQALPLPATGKNKVLGKKGDVTELTCTASQKSIQFMKNSNDIK 60

QY 61 ILNGSFLTKGPSKLANDRADSRSLMDQNPPLIKNLKIEDSDTYICEVEDQKEEYOL 120
Db 61 ILNGSFLTKGPSKLANDRADSRSLMDQNPPLIKNLKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSDBTHLQGSGLTTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGLTANSDBTHLQGSGLTTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCVTLONOKKVEFKIDIVLAFQKASSIYVKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
Db 181 TWTCVTLONOKKVEFKIDIVLAFQKASSIYVKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSSKSWITFDLKNKEVSKRVTPDKLQMGKULPLHLTLPOALPOYAGSGLTLA 300
Db 241 QAERASSSSKSWITFDLKNKEVSKRVTPDKLQMGKULPLHLTLPOALPOYAGSGLTLA 300
QY 301 LEAKTGKLGHOEVNLYVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPYWV 360
Db 301 LEAKTGKLGHOEVNLYVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPYWV 360
QY 361 INPEAGMWOCILSDSGQVLESNIKVLPTWSTPVHPRASALPAPPTGSALPDPOTASALP 420
Db 361 INPEAGMWOCILSDSGQVLESNIKVLPTWSTPVH-----ADPQ----- 399
QY 421 DPPASALPALAVISFLGLGL 443
Db 400 -----LCYILDALFLYGLVL 415
RESULT 7
PCT-US95-00454-5
Sequence 5, Application PC/TUS9500454
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 462 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-00454-5

Query Match 87.7%; Score 2040; DB 5; Length 462;
 Best Local Similarity 91.4%; Pred. No. 2e-162;
 Matches 405; Conservative 1; Mismatches 9; Indels 28; Gaps 2;

```

QY 1 MNRGVPRHLLVQLALLPAATQGNKVLGKGGDTVELTCTASQKKSIOFHWKNSNOIK 60
DB 1 MNRGVPRHLLVQLALLPAATQGNKVLGKGGDTVELTCTASQKKSIOFHWKNSNOIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIKKLTIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIKKLTIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGSGLTLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGSGLTLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTCVLQONKKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWCTCVLQONKKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
QY 241 QAEBASSSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGLTLA 300
DB 241 QAEBASSSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGLTLA 300
QY 301 LEAKTGKLGHEVNLVWMRATOLQKNLTCEVWGPTSPKMLSLKENKEAVSKREKPVW 360
DB 301 LEAKTGKLGHEVNLVWMRATOLQKNLTCEVWGPTSPKMLSLKENKEAVSKREKPVW 360
QY 361 LNPEAGMWQCLLSDSGVLLSNIKVLPTWSTPVHAPRASPALPPTGSALDPQTASALP 420
DB 361 LNPEAGMWQCLLSDSGVLLSNIKVLPTWSTPVHAPRASPALPPTGSALDPQTASALP 420
QY 421 DPBAASALPALAVISFLGLGL 443
DB 400 -----LCYTLDAFLFLXGIVL 415

```

RESULT 8

US-08-417-495-4
 Sequence 4, Application US/08417495
 Patent No. 5843728

GENERAL INFORMATION:
 APPLICANT: Seed, Brian et al.
 TITLE OF INVENTION: Redirection of Cellular Immunity by Chimerae
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 502 or 555X
 OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 SOFTWARE: Wordperfect (Version 5.0)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08417,495

FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/203,866
 FILING DATE:
 APPLICATION NUMBER: US/07/847,566
 FILING DATE:

APPLICATION NUMBER: 07/665,961
 FILING DATE: March 7, 1991
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 00786/119002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 575 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-417-495-4

Query Match 87.7%; Score 2038.5; DB 2; Length 575;
 Best Local Similarity 92.1%; Pred. No. 3.7e-162;
 Matches 406; Conservative 4; Mismatches 8; Indels 23; Gaps 3;

```

QY 1 MNRGVPRHLLVQLALLPAATQGNKVLGKGGDTVELTCTASQKKSIOFHWKNSNOIK 60
DB 1 MNRGVPRHLLVQLALLPAATQGNKVLGKGGDTVELTCTASQKKSIOFHWKNSNOIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIKKLTIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIKKLTIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGSGLTLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGSGLTLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTCVLQONKKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWCTCVLQONKKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
QY 241 QAEBASSSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGLTLA 300
DB 241 QAEBASSSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGLTLA 300
QY 301 LEAKTGKLGHEVNLVWMRATOLQKNLTCEVWGPTSPKMLSLKENKEAVSKREKPVW 360
DB 301 LEAKTGKLGHEVNLVWMRATOLQKNLTCEVWGPTSPKMLSLKENKEAVSKREKPVW 360
QY 361 LNPEAGMWQCLLSDSGVLLSNIKVLPTWSTPVHAPRASPALPPTGSALDPQTASALP 420
DB 361 LNPEAGMWQCLLSDSGVLLSNIKVLPTWSTPVHAPRASPALPPTGSALDPQTASALP 420
QY 402 PAPPTGSALDPQTASALPDP 422
DB 421 RAKFRSA---ETAAVLQDP 437

```

RESULT 9

US-08-284-391B-4
 Sequence 4, Application US/08284391B
 Patent No. 5851828

GENERAL INFORMATION:
 APPLICANT: Seed, Brian
 TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
 NUMBER OF SEQUENCES: 53
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Clark & Elbing LLP
 STREET: 176 Federal Street
 CITY: Boston
 STATE: MA

COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Eibling, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-4

Query Match 87.7% Score 2038.5; DB 2; Length 575;
Best Local Similarity 92.1% Pred. No. 3.7e-162;

Matches 406; Conservative 4; Mismatches 8; Indels 23; Gaps 3;

QY 1 MNRGVFPHLLLVLTALPPATOGNKVVLGKKGDTVELTCTASOKSIOFHMKNSNOIK 60
DB 1 MNRGVFPHLLLVLTALPPATOGNKVVLGKKGDTVELTCTASOKSIOFHMKNSNOIK 60
QY 61 ILNGSGFLTKGPSKLNDRADSRSLMDQGNPPLIKLKIEDSDTYICEVEDOKEEVOL 120
DB 61 ILNGSGFLTKGPSKLNDRADSRSLMDQGNPPLIKLKIEDSDTYICEVEDOKEEVOL 120
QY 121 LVFGLTANSDBTHLLOGSILTLTLESPPGSSPVOCRSRPGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDBTHLLOGSILTLTLESPPGSSPVOCRSRPGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTVLOQKQKVEFKIDIVLAFQKASSIYKKGEQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWCTVLOQKQKVEFKIDIVLAFQKASSIYKKGEQVEFSPLAFTVEKLTGSGELMW 240
QY 241 QAEKASSSSKSWITFDLKNKEVSQKRVTPDKLQMGKPLHLTLQALPQVYAGSNTLTA 300
DB 241 QAEKASSSSKSWITFDLKNKEVSQKRVTPDKLQMGKPLHLTLQALPQVYAGSNTLTA 300
QY 301 LEAKGKJLHQLQEVNLLVMRATOLQKULTCVWGPTSPKMLSLKLNKEKAKSKREKPVAV 360
DB 301 LEAKGKJLHQLQEVNLLVMRATOLQKULTCVWGPTSPKMLSLKLNKEKAKSKREKPVAV 360
QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPYVSTPVH--PR-----ASAL 401
DB 361 LNPEAGMOCCLSDSGOVLLESNIKVLPYVSTPVH--PR-----ASAL 401
QY 402 PAPPTGSALPDQPTASALPDP 422
DB 422 PAPPTGSALPDQPTASALPDP 422
QY 421 RAKFRSA---ETANLQDP 437
DB 421 RAKFRSA---ETANLQDP 437

RESULT 10

US-09-218-950-4
Sequence 4, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Eibling LLP
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/847,566
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Eibling, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-4

Query Match 87.7% Score 2038.5; DB 3; Length 575;
Best Local Similarity 92.1% Pred. No. 3.7e-162;

Matches 406; Conservative 4; Mismatches 8; Indels 23; Gaps 3;

QY 1 MNRGVFPHLLLVLTALPPATOGNKVVLGKKGDTVELTCTASOKSIOFHMKNSNOIK 60
DB 1 MNRGVFPHLLLVLTALPPATOGNKVVLGKKGDTVELTCTASOKSIOFHMKNSNOIK 60
QY 61 ILNGSGFLTKGPSKLNDRADSRSLMDQGNPPLIKLKIEDSDTYICEVEDOKEEVOL 120
DB 61 ILNGSGFLTKGPSKLNDRADSRSLMDQGNPPLIKLKIEDSDTYICEVEDOKEEVOL 120
QY 121 LVFGLTANSDBTHLLOGSILTLTLESPPGSSPVOCRSRPGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDBTHLLOGSILTLTLESPPGSSPVOCRSRPGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTVLOQKQKVEFKIDIVLAFQKASSIYKKGEQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWCTVLOQKQKVEFKIDIVLAFQKASSIYKKGEQVEFSPLAFTVEKLTGSGELMW 240

QY 241 QABASSSSKSWITFDLKNKEVSVKRVTOPDKLQMGKKLPLHLTLPOLPYAGSGNLTLLA 300
DB 241 QABASSSSKSWITFDLKNKEVSVKRVTOPDKLQMGKKLPLHLTLPOLPYAGSGNLTLLA 300
QY 301 LEATGKGLHOEVNLVVMRATQLOKNTCEVWGPTSPKMLSLKLENKAKVSKREKPVW 360
DB 301 LEATGKGLHOEVNLVVMRATQLOKNTCEVWGPTSPKMLSLKLENKAKVSKREKPVW 360
QY 361 LNPEAGMWQCLLSDSGQVLESNIKVLPTWSTPVH--PR-----ASAL 401
DB 361 LNPEAGMWQCLLSDSGQVLESNIKVLPTWSTPVHADPKLCYLLDGLIFVGYITLALYL 420
QY 402 PAPPTGSALPDPQTASALPDP 422
DB 421 RAKFSRSA-----ETANLQDP 437

RESULT 11

PCT-US92-01785-4
Sequence 4, Application PC/TUS9201785
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-01785-4

Query Match 87.7%; Score 2038.5; DB 5; Length 575;
Best Local Similarity 92.1%; Pred. No. 3.7e-162;
Matches 406; Conservative 4; Mismatches 8; Indels 23; Gaps 3;

QY 1 MNRGVFPHLLVQLALPAATQGNKVLAGKGDVTELTCTASOKSIOFHMKNNOIK 60
DB 1 MNRGVFPHLLVQLALPAATQGNKVLAGKGDVTELTCTASOKSIOFHMKNNOIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIKLKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIKLKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSDFTHLLQGGSLTLTLESPPGSSPSVOCRSFRKNIOGKTLVSQLELQDSG 180

DB 121 LVFGLTANSDFTHLLQGGSLTLTLESPPGSSPSVOCRSFRKNIOGKTLVSQLELQDSG 180
QY 181 TWCTVLOKOKKVEFKDIDIVLAFOKASSIVYKKEGQVFSPLAFTVETLGGSGELMW 240
DB 181 TWCTVLOKOKKVEFKDIDIVLAFOKASSIVYKKEGQVFSPLAFTVETLGGSGELMW 240
QY 241 QABASSSSKSWITFDLKNKEVSVKRVTOPDKLQMGKKLPLHLTLPOLPYAGSGNLTLLA 300
DB 241 QABASSSSKSWITFDLKNKEVSVKRVTOPDKLQMGKKLPLHLTLPOLPYAGSGNLTLLA 300
QY 301 LEATGKGLHOEVNLVVMRATQLOKNTCEVWGPTSPKMLSLKLENKAVSKREKPVW 360
DB 301 LEATGKGLHOEVNLVVMRATQLOKNTCEVWGPTSPKMLSLKLENKAVSKREKPVW 360
QY 361 LNPEAGMWQCLLSDSGQVLESNIKVLPTWSTPVH--PR-----ASAL 401
DB 361 LNPEAGMWQCLLSDSGQVLESNIKVLPTWSTPVHADPKLCYLLDGLIFVGYITLALYL 420
QY 402 PAPPTGSALPDPQTASALPDP 422
DB 421 RAKFSRSA-----ETANLQDP 437

RESULT 12

PCT-US95-00454-4
Sequence 4, Application PC/TUS9500454
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00454-4

```

Query Match      87.7%; Score 2038.5; DB 5; Length 575;
Best Local Similarity 92.1%; Pred. No. 3.7e-162;
Matches 406; Conservative 4; Mismatches 8; Indels 23; Gaps 3;

QY 1 NMRGVPFPHLLVLQALLPAATQGNKVVLGKKGDVETLTASQKKSIOFHMKNSNOJK 60
DB 1 NMRGVPFPHLLVLQALLPAATQGNKVVLGKKGDVETLTASQKKSIOFHMKNSNOJK 60
QY 61 IINGGGSFLTKGPSKLNDRADSRSLMDQGNPPLIIKNIKIEDSDTYICEVDQKEEYOL 120
DB 61 IINGGGSFLTKGPSKLNDRADSRSLMDQGNPPLIIKNIKIEDSDTYICEVDQKEEYOL 120
QY 121 LVFGLTANS DTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANS DTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 121 LVFGLTANS DTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANS DTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLQNOQKVEFKIDIVLAFQKASSIYVKKGEQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLQNOQKVEFKIDIVLAFQKASSIYVKKGEQVEFSPLAFTVEKLTGSGELMW 240
QY 241 QAEKASSSSKSWITFDLKNKEVSVKRVTPDKLQMGKKLPLHLTLPQALPQVAGSGLTLA 300
DB 241 QAEKASSSSKSWITFDLKNKEVSVKRVTPDKLQMGKKLPLHLTLPQALPQVAGSGLTLA 300
QY 301 LEAKTGKLEHGVNLYVMRATOLQKNLTCEWGPSPKMLSLKLEKKAQVSKREKPVAV 360
DB 301 LEAKTGKLEHGVNLYVMRATOLQKNLTCEWGPSPKMLSLKLEKKAQVSKREKPVAV 360
QY 361 INPEAGMOCCLSDSGVLLBSNIVKLPWTWSTPVH--PR-----ASAL 401
DB 361 INPEAGMOCCLSDSGVLLBSNIVKLPWTWSTPVHADPKLQVLLDGLFTYGIITALLV 420
QY 402 PAPPGSALPDPTQASALPD 422
DB 421 RAKFSRSA---ETANLQDP 437

RESULT 13
US-08-328-500-9
; Sequence 9, Application US/08328500
; Patent No. 6673896
; GENERAL INFORMATION:
; APPLICANT: Maddon, Paul J.
; APPLICANT: Axel, Richard W.
; APPLICANT: Sweet, Richard W.
; APPLICANT: Arthos, James
; TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; FILING DATE:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/24577-CY
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 457 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-328-500-9

Query Match      87.7%; Score 2038; DB 4; Length 457;
Best Local Similarity 89.4%; Pred. No. 2.9e-162;
Matches 406; Conservative 2; Mismatches 16; Indels 30; Gaps 2;

QY 1 NMRGVPFPHLLVLQALLPAATQGNKVVLGKKGDVETLTASQKKSIOFHMKNSNOJK 60
DB 1 NMRGVPFPHLLVLQALLPAATQGNKVVLGKKGDVETLTASQKKSIOFHMKNSNOJK 60
QY 61 IINGGGSFLTKGPSKLNDRADSRSLMDQGNPPLIIKNIKIEDSDTYICEVDQKEEYOL 120
DB 61 IINGGGSFLTKGPSKLNDRADSRSLMDQGNPPLIIKNIKIEDSDTYICEVDQKEEYOL 120
QY 121 LVFGLTANS DTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANS DTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 121 LVFGLTANS DTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANS DTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLQNOQKVEFKIDIVLAFQKASSIYVKKGEQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLQNOQKVEFKIDIVLAFQKASSIYVKKGEQVEFSPLAFTVEKLTGSGELMW 240
QY 241 QAEKASSSSKSWITFDLKNKEVSVKRVTPDKLQMGKKLPLHLTLPQALPQVAGSGLTLA 300
DB 241 QAEKASSSSKSWITFDLKNKEVSVKRVTPDKLQMGKKLPLHLTLPQALPQVAGSGLTLA 300
QY 301 LEAKTGKLEHGVNLYVMRATOLQKNLTCEWGPSPKMLSLKLEKKAQVSKREKPVAV 360
DB 301 LEAKTGKLEHGVNLYVMRATOLQKNLTCEWGPSPKMLSLKLEKKAQVSKREKPVAV 360
QY 361 INPEAGMOCCLSDSGVLLBSNIVKLPWTWSTPVHPRASALPAPPTGSALPDPTQASALP 420
DB 361 INPEAGMOCCLSDSGVLLBSNIVKLPWTWSTPVHPRASALPAPPTGSALPDPTQASALP 420
QY 421 DPPASALPALAVISFLGLGLV-ACVLAATR 453
DB 397 -----MALIVLGVAGVGLLFLGLGIFCVGRGHR 425

RESULT 14
US-08-284-391B-29
; Sequence 29, Application US/08284391B
; Patent No. 5851828
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: PastsEO for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 02-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/195,395

```

FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Eibling, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-29

Query Match 87.6%; Score 2037; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 2.8e-162;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRLHLVQLALLPAATQGNKRVVLGKKGDVTELCTASOKSIOFHWKNSNOIK 60
DB 1 MNRGVPFRLHLVQLALLPAATQGNKRVVLGKKGDVTELCTASOKSIOFHWKNSNOIK 60
QY 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNFPIIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNFPIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGQSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGQSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCVTLOQNKVEFKIDIVVLAFOKASSIVYKKEGEQVEFSPLATVEKLTGSGGLMW 240
DB 181 TWTCVTLOQNKVEFKIDIVVLAFOKASSIVYKKEGEQVEFSPLATVEKLTGSGGLMW 240
QY 241 QAEBAASSKSWITFDLKNKEVSVKRVTDQPKLQMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
DB 241 QAEBAASSKSWITFDLKNKEVSVKRVTDQPKLQMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
QY 301 LEAKTGKLGHOENVLVVWRATQKNIJCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360
DB 301 LEAKTGKLGHOENVLVVWRATQKNIJCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360
QY 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTPVH 395
DB 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTPVH 395

RESULT 15

US-09-218-950-29
Sequence 29, Application US/09218950

Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Maideemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Eibling LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA

ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Eibling, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-29

Query Match 87.6%; Score 2037; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 2.8e-162;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRLHLVQLALLPAATQGNKRVVLGKKGDVTELCTASOKSIOFHWKNSNOIK 60
DB 1 MNRGVPFRLHLVQLALLPAATQGNKRVVLGKKGDVTELCTASOKSIOFHWKNSNOIK 60
QY 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNFPIIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNFPIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGQSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGQSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCVTLOQNKVEFKIDIVVLAFOKASSIVYKKEGEQVEFSPLATVEKLTGSGGLMW 240
DB 181 TWTCVTLOQNKVEFKIDIVVLAFOKASSIVYKKEGEQVEFSPLATVEKLTGSGGLMW 240
QY 241 QAEBAASSKSWITFDLKNKEVSVKRVTDQPKLQMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
DB 241 QAEBAASSKSWITFDLKNKEVSVKRVTDQPKLQMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
QY 301 LEAKTGKLGHOENVLVVWRATQKNIJCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360
DB 301 LEAKTGKLGHOENVLVVWRATQKNIJCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360
QY 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTPVH 395
DB 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTPVH 395

RESULT 16

US-08-417-495-6
Sequence 6, Application US/08417495
Patent No. 5843728

```

GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Redirection of Cellular Immunity by
TITLE OF INVENTION: Receptor
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,495
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/203,866
FILING DATE:
APPLICATION NUMBER: US/07/847,566
FILING DATE:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-417-495-6

Query Match      87.6%; Score 2037; DB 2; Length 532;
Best Local Similarity 100.0%; Pred. No. 4,4e-162;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVLTQALLPAAATGKNVVLGKKGDTVELTCTAQQKSIQPHMKNNSQIK 60
DB 1 MNRGVPFRHLLVLTQALLPAAATGKNVVLGKKGDTVELTCTAQQKSIQPHMKNNSQIK 60
QY 61 ILNGSGFLTGKPSKLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVQL 120
DB 61 ILNGSGFLTGKPSKLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVQL 120
QY 121 LVFGLTANSDFHLLQGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSVQLELQDSG 180
DB 121 LVFGLTANSDFHLLQGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSVQLELQDSG 180
QY 181 TWTCTVLONOKKVEKIDIVVLAFOKASSIYKKEGEQEVSRFLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLONOKKVEKIDIVVLAFOKASSIYKKEGEQEVSRFLAFTVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTODPKLQMGKLLPLHLTLPQALPOYAGSGLNTILA 300
DB 241 QAERASSSKSWITFDLKNKEVSVKRVTODPKLQMGKLLPLHLTLPQALPOYAGSGLNTILA 300
QY 301 LEAKTGKGLHGVNLVVMRATOLQKNLTCEVWGFTSPKMLSLIKENKEAKVSKREKPVVW 360
DB 301 LEAKTGKGLHGVNLVVMRATOLQKNLTCEVWGFTSPKMLSLIKENKEAKVSKREKPVVW 360
```

```

QY 361 LNPEAGMWQCLSDSGVLTLESNIKVLPTWSTPVA 395
DB 361 LNPEAGMWQCLSDSGVLTLESNIKVLPTWSTPVA 395

RESULT 17
US-08-284-391B-6
Sequence 6, Application US/08284391B
Patent No. 5851828
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSER: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-6

Query Match      87.6%; Score 2037; DB 2; Length 532;
Best Local Similarity 100.0%; Pred. No. 4,4e-162;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVLTQALLPAAATGKNVVLGKKGDTVELTCTAQQKSIQPHMKNNSQIK 60
DB 1 MNRGVPFRHLLVLTQALLPAAATGKNVVLGKKGDTVELTCTAQQKSIQPHMKNNSQIK 60
QY 61 ILNGSGFLTGKPSKLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVQL 120
DB 61 ILNGSGFLTGKPSKLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVQL 120
QY 121 LVFGLTANSDFHLLQGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSVQLELQDSG 180
DB 121 LVFGLTANSDFHLLQGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSVQLELQDSG 180
```

Qy 181 TWCTVLONOKKVEFKIDIVLAFQKASSIVYKKEGQVFSPLAFTVEKLTGSGELMW 240
Db 181 TWCTVLONOKKVEFKIDIVLAFQKASSIVYKKEGQVFSPLAFTVEKLTGSGELMW 240
Qy 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKKLPLHLTPQALPOYAGSGNLTLLA 300
Db 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKKLPLHLTPQALPOYAGSGNLTLLA 300
Qy 301 LEAKTGKLGHEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPYWV 360
Db 301 LEAKTGKLGHEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPYWV 360
Qy 361 LNPEAGMWQCLSDSGQVLLSESNIKVLPWTSTPVH 395
Db 361 LNPEAGMWQCLSDSGQVLLSESNIKVLPWTSTPVH 395

RESULT 18
US-09-218-950-6
Sequence 6, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218.950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284.391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195.395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847.566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665.961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35.238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-6

Query Match 87.6%; Score 2037; DB 3; Length 532;
Best Local Similarity 100.0%; Pred. No. 4,4e-162;

Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MNRGVPFRHLVLYOLALPAATQGNKVYLGKKDFTVELCTASQOKSIQFHMKNNSQIK 60
Db 1 MNRGVPFRHLVLYOLALPAATQGNKVYLGKKDFTVELCTASQOKSIQFHMKNNSQIK 60
Qy 61 ILNQGSLFKTKGSPKLNDRASRRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEVQL 120
Db 61 ILNQGSLFKTKGSPKLNDRASRRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEVQL 120
Qy 121 LVFGILTANSTHLLQGOQLTLTLESPGSSPSVQCSPPRKNIOGGKTLTSSQLTLDG 180
Db 121 LVFGILTANSTHLLQGOQLTLTLESPGSSPSVQCSPPRKNIOGGKTLTSSQLTLDG 180
Qy 181 TWCTVLONOKKVEFKIDIVLAFQKASSIVYKKEGQVFSPLAFTVEKLTGSGELMW 240
Db 181 TWCTVLONOKKVEFKIDIVLAFQKASSIVYKKEGQVFSPLAFTVEKLTGSGELMW 240
Qy 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKKLPLHLTPQALPOYAGSGNLTLLA 300
Db 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKKLPLHLTPQALPOYAGSGNLTLLA 300
Qy 301 LEAKTGKLGHEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPYWV 360
Db 301 LEAKTGKLGHEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPYWV 360
Qy 361 LNPEAGMWQCLSDSGQVLLSESNIKVLPWTSTPVH 395
Db 361 LNPEAGMWQCLSDSGQVLLSESNIKVLPWTSTPVH 395

RESULT 19
PCT-US92-01785-6
Sequence 6, Application PC/TUS9201785
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-01785-6

Query Match 87.6%; Score 2037; DB 5; Length 532;
Best Local Similarity 100.0%; Pred. No. 4.4e-162;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPPAATQGNKVLGKGGDTVELTCTASQKSIQFHWKXSNQIK 60
DB 1 MNRGVPFRHLVLVQLALPPAATQGNKVLGKGGDTVELTCTASQKSIQFHWKXSNQIK 60
QY 61 ILNGSGFLTKGPSKLNDRADSRSLMDQGNPPLIIKLIKIDSDTYICEVEDQKEEYQL 120
DB 61 ILNGSGFLTKGPSKLNDRADSRSLMDQGNPPLIIKLIKIDSDTYICEVEDQKEEYQL 120
QY 121 LVFGLTANSDDTHLQGGSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDDTHLQGGSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTVLOKQKVEFKIDIVVLAFOKASSIYKKEGEQVEFSFPLAFTVEKLTSGGELMW 240
DB 181 TWCTVLOKQKVEFKIDIVVLAFOKASSIYKKEGEQVEFSFPLAFTVEKLTSGGELMW 240
QY 241 QAERASSSKSWITFPDLKNKEVSVKRVTDPKLQMGKULPLHLTLPQALPOYAGSGLTLIA 300
DB 241 QAERASSSKSWITFPDLKNKEVSVKRVTDPKLQMGKULPLHLTLPQALPOYAGSGLTLIA 300
QY 301 LEAKTGKLGHOEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKENKEAKVSKREKPYWV 360
DB 301 LEAKTGKLGHOEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKENKEAKVSKREKPYWV 360
QY 361 LNPEAGMWQCLSDSGQVLLBSNIKVLPWTSTPVH 395
DB 361 LNPEAGMWQCLSDSGQVLLBSNIKVLPWTSTPVH 395

RESULT 20
PCT-US95-00454-6
; Sequence 6, Application PC/TUS9500454
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
; TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
; TITLE OF INVENTION: Bearing Cells
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 558X
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/00454
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/847,566
; FILING DATE: March 6, 1992
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/665,961
; FILING DATE: March 7, 1991
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-00454-6

Query Match 87.6%; Score 2037; DB 5; Length 532;
Best Local Similarity 100.0%; Pred. No. 4.4e-162;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPPAATQGNKVLGKGGDTVELTCTASQKSIQFHWKXSNQIK 60
DB 1 MNRGVPFRHLVLVQLALPPAATQGNKVLGKGGDTVELTCTASQKSIQFHWKXSNQIK 60
QY 61 ILNGSGFLTKGPSKLNDRADSRSLMDQGNPPLIIKLIKIDSDTYICEVEDQKEEYQL 120
DB 61 ILNGSGFLTKGPSKLNDRADSRSLMDQGNPPLIIKLIKIDSDTYICEVEDQKEEYQL 120
QY 121 LVFGLTANSDDTHLQGGSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDDTHLQGGSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTVLOKQKVEFKIDIVVLAFOKASSIYKKEGEQVEFSFPLAFTVEKLTSGGELMW 240
DB 181 TWCTVLOKQKVEFKIDIVVLAFOKASSIYKKEGEQVEFSFPLAFTVEKLTSGGELMW 240
QY 241 QAERASSSKSWITFPDLKNKEVSVKRVTDPKLQMGKULPLHLTLPQALPOYAGSGLTLIA 300
DB 241 QAERASSSKSWITFPDLKNKEVSVKRVTDPKLQMGKULPLHLTLPQALPOYAGSGLTLIA 300
QY 301 LEAKTGKLGHOEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKENKEAKVSKREKPYWV 360
DB 301 LEAKTGKLGHOEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKENKEAKVSKREKPYWV 360
QY 361 LNPEAGMWQCLSDSGQVLLBSNIKVLPWTSTPVH 395
DB 361 LNPEAGMWQCLSDSGQVLLBSNIKVLPWTSTPVH 395

RESULT 21
US-09-039-555B-15
; Sequence 15, Application US/09039555B
; Patent No. 6033856
; GENERAL INFORMATION:
; APPLICANT: Koerner, Kathrin
; APPLICANT: Mueller, Rolf
; TITLE OF INVENTION: Sadiacek, Hans-Harald
; TITLE OF INVENTION: PROMOTER OF THE CDC25B GENE, ITS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Hardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,555B
; FILING DATE: 16-MAR-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 19710643.9
; FILING DATE: 14-MAR-1997

ATTORNEY/AGENT INFORMATION:

NAME: Bent, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 016779/0131

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 458 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-039-555B-15

Query Match 87.4%; Score 2032; DB 3; Length 458;

Best Local Similarity 89.2%; Pred. No. 9,1e-162;

Matches 405; Conservative 2; Mismatches 17; Indels 30; Gaps 2;

```
QY 1 MNRGVPFRHLLVQLALPPATQGNKVVLGKGDVVELTCTASQKSIQPHMKNQIK 60
DB 1 MNRGVPFRHLLVQLALPPATQGNKVVLGKGDVVELTCTASQKSIQPHMKNQIK 60
QY 61 ILNGSGFLTKGPKSLNDRADSRSLMDQGNFPLIKNLKIETSDTYICEVEDQKEEVOL 120
DB 61 ILNGSGFLTKGPKSLNDRADSRSLMDQGNFPLIKNLKIETSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
DB 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
QY 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
DB 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
QY 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
DB 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
QY 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
DB 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
QY 241 QABRASSKSWITFDLKNKEVSVKRVTPDPLQMGKKLPLHLTLPQALPYAGSGNLTLA 300
DB 241 QABRASSKSWITFDLKNKEVSVKRVTPDPLQMGKKLPLHLTLPQALPYAGSGNLTLA 300
QY 301 LEATGKLGHOEVLVVMRATQLOKNLTCEVWGPTSPKLMSTLKENKAVSKREKVPVW 360
DB 301 LEATGKLGHOEVLVVMRATQLOKNLTCEVWGPTSPKLMSTLKENKAVSKREKVPVW 360
QY 361 LNPBAGMWQCLSDSGOVLESNIKVLPTWSTVPHPRASALPAPPTGSALPDQTASALP 420
DB 361 LNPBAGMWQCLSDSGOVLESNIKVLPTWSTVPHPRASALPAPPTGSALPDQTASALP 420
QY 421 DPPAASALPALAVISFLGLGLGV-ACVLARTR 453
DB 397 -----MALIVLGGVAGLLIFGLGIFPCVRCRRH 425
```

RESULT 22

US-09-517-605-3

Sequence 3, Application US/09517605

Patent No. 6391567

GENERAL INFORMATION:

APPLICANT: Littman, Dan R.

APPLICANT: Kwon, Douglas S.

APPLICANT: van Kooyk, Yvette

APPLICANT: Geiltebeck, Theo

TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO

FILE REFERENCE: 1049-1-017

CURRENT APPLICATION NUMBER: US/09/517, 605

CURRENT FILING DATE: 2000-03-02

NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 458

TYPE: PRT

ORGANISM: Homo sapiens

US-09-517-605-3

Query Match 87.1%; Score 2024; DB 4; Length 458;

Best Local Similarity 89.0%; Pred. No. 4,3e-161;

Matches 404; Conservative 2; Mismatches 18; Indels 30; Gaps 2;

```
QY 1 MNRGVPFRHLLVQLALPPATQGNKVVLGKGDVVELTCTASQKSIQPHMKNQIK 60
DB 1 MNRGVPFRHLLVQLALPPATQGNKVVLGKGDVVELTCTASQKSIQPHMKNQIK 60
QY 61 ILNGSGFLTKGPKSLNDRADSRSLMDQGNFPLIKNLKIETSDTYICEVEDQKEEVOL 120
DB 61 ILNGSGFLTKGPKSLNDRADSRSLMDQGNFPLIKNLKIETSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
DB 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
QY 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
DB 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
QY 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
DB 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
QY 241 QABRASSKSWITFDLKNKEVSVKRVTPDPLQMGKKLPLHLTLPQALPYAGSGNLTLA 300
DB 241 QABRASSKSWITFDLKNKEVSVKRVTPDPLQMGKKLPLHLTLPQALPYAGSGNLTLA 300
QY 301 LEATGKLGHOEVLVVMRATQLOKNLTCEVWGPTSPKLMSTLKENKAVSKREKVPVW 360
DB 301 LEATGKLGHOEVLVVMRATQLOKNLTCEVWGPTSPKLMSTLKENKAVSKREKVPVW 360
QY 361 LNPBAGMWQCLSDSGOVLESNIKVLPTWSTVPHPRASALPAPPTGSALPDQTASALP 420
DB 361 LNPBAGMWQCLSDSGOVLESNIKVLPTWSTVPHPRASALPAPPTGSALPDQTASALP 420
QY 421 DPPAASALPALAVISFLGLGLGV-ACVLARTR 453
DB 397 -----MALIVLGGVAGLLIFGLGIFPCVRCRRH 425
```

RESULT 23

US-08-236-311-1

Sequence 1, Application US/08236311

Patent No. 5565335

GENERAL INFORMATION:

APPLICANT: Capon, Daniel J.

APPLICANT: Gregory, Timothy J.

TITLE OF INVENTION: Adhesion Variants

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/236,311

FILING DATE: 02-MAY-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/936190

FILING DATE: 26-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/842777

FILING DATE: 18-FEB-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/250785

```

; FILING DATE: 28-SEP-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Habak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 444P1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1996
; TELEFAX: 415/952-9881
; TELEEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-236-311-1

```

```

Query Match      86.8%; Score 2017; DB 1; Length 402;
Best Local Similarity 99.7%; Pred. No. 1,4e-160;
Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPFRHLILVQLALPPAATQGNKVVLGKKGDVVELTCTASQKSIQFHMKNNOIK 60
   |||||
DB 1 MNRGVPFRHLILVQLALPPAATQGNKVVLGKKGDVVELTCTASQKSIQFHMKNNOIK 60
QY 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNFPLIIKLIKIEDSPDYICEVEDQKEEYOL 120
   |||||
DB 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNFPLIIKLIKIEDSPDYICEVEDQKEEYOL 120
QY 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNFPLIIKLIKIEDSPDYICEVEDQKEEYOL 120
   |||||
DB 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNFPLIIKLIKIEDSPDYICEVEDQKEEYOL 120
QY 121 LVFGLTANSDBTHLQGGSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
   |||||
DB 121 LVFGLTANSDBTHLQGGSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTCTVLQNKQKVEFKIDIVLAFQKASSIYKKEGEVEFSFPLAFTVEKLTSGGELMW 240
   |||||
DB 181 TWCTCTVLQNKQKVEFKIDIVLAFQKASSIYKKEGEVEFSFPLAFTVEKLTSGGELMW 240
QY 241 QAEKSSSSKSWITPDLKKEVSVKRVTPDKLQMGKCLPLHLTPALPOVAGSNNLTIA 300
   |||||
DB 241 QAEKSSSSKSWITPDLKKEVSVKRVTPDKLQMGKCLPLHLTPALPOVAGSNNLTIA 300
QY 301 LEAKTGKLGHEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPYVW 360
   |||||
DB 301 LEAKTGKLGHEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPYVW 360
QY 361 LNPEAGMMQCLSDSGQVLLSNIKVLPTWSTP 393
   |||||
DB 361 LNPEAGMMQCLSDSGQVLLSNIKVLPTWSTP 393

```

```

RESULT 24
US-08-457-918-1
; Sequence 1, Application US/08457918
; Patent No. 6117655
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; APPLICANT: Gregory, Timothy J.
; TITLE OF INVENTION: Adhesion Variance
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,918
; FILING DATE: 1-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/236311
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/936190
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/842777
; FILING DATE: 18-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/250785
; FILING DATE: 28-SEP-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0444P1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8228
; TELEFAX: 415/952-9881
; TELEEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-457-918-1

```

```

Query Match      86.8%; Score 2017; DB 3; Length 402;
Best Local Similarity 99.7%; Pred. No. 1,4e-160;
Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPFRHLILVQLALPPAATQGNKVVLGKKGDVVELTCTASQKSIQFHMKNNOIK 60
   |||||
DB 1 MNRGVPFRHLILVQLALPPAATQGNKVVLGKKGDVVELTCTASQKSIQFHMKNNOIK 60
QY 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNFPLIIKLIKIEDSPDYICEVEDQKEEYOL 120
   |||||
DB 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNFPLIIKLIKIEDSPDYICEVEDQKEEYOL 120
QY 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNFPLIIKLIKIEDSPDYICEVEDQKEEYOL 120
   |||||
DB 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNFPLIIKLIKIEDSPDYICEVEDQKEEYOL 120
QY 121 LVFGLTANSDBTHLQGGSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
   |||||
DB 121 LVFGLTANSDBTHLQGGSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTCTVLQNKQKVEFKIDIVLAFQKASSIYKKEGEVEFSFPLAFTVEKLTSGGELMW 240
   |||||
DB 181 TWCTCTVLQNKQKVEFKIDIVLAFQKASSIYKKEGEVEFSFPLAFTVEKLTSGGELMW 240
QY 241 QAEKSSSSKSWITPDLKKEVSVKRVTPDKLQMGKCLPLHLTPALPOVAGSNNLTIA 300
   |||||
DB 241 QAEKSSSSKSWITPDLKKEVSVKRVTPDKLQMGKCLPLHLTPALPOVAGSNNLTIA 300
QY 301 LEAKTGKLGHEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPYVW 360
   |||||
DB 301 LEAKTGKLGHEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPYVW 360
QY 361 LNPEAGMMQCLSDSGQVLLSNIKVLPTWSTP 393
   |||||
DB 361 LNPEAGMMQCLSDSGQVLLSNIKVLPTWSTP 393

```

```

RESULT 25
5223394-7
; Patent No. 5223394
; APPLICANT: WALLNER, BARBARA
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
; LYMPHOCTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL

```

LINKAGE SIGNAL SEQUENCE
 NUMBER OF SEQUENCES: 12
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/335,688
 FILING DATE: 10-APR-1989
 SEQ ID NO: 7
 LENGTH: 458
 5223394-7

Query Match 86.3%; Score 2006; DB 6; Length 458;
 Best Local Similarity 88.3%; Pred. No. 1.4e-159;
 Matches 401; Conservative 2; Mismatches 21; Indels 30; Gaps 2;

QY 1 MNRGVPPRHLLVQLALPPAATOGKVVYLGKGDVVELTCTASOKKSIOFHMKNNOIK 60
 DB 1 MNRGVPPRHLLVQLALPPAATOGKVVYLGKGDVVELTCTASOKKSIOFHMKNNOIK 60
 QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDYICEVEDQKEEVOL 120
 DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDYICEVEDQKEEVOL 120
 QY 121 LVFGLTANSPTHLLQGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
 DB 121 LVFGLTANSPTHLLQGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
 QY 181 TWCTVLQONQKVEFKIDIVLAFOKASSIYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
 DB 181 TWCTVLQONQKVEFKIDIVLAFOKASSIYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
 QY 241 QAEKASSKSWITFDLKNKEVSVKRVTQDPKLOMGKPLHLTLPOALPOYAGSGNLTLLA 300
 DB 241 QAEKASSKSWITFDLKNKEVSVKRVTQDPKLOMGKPLHLTLPOALPOYAGSGNLTLLA 300
 QY 301 LEAKTGKLGHOEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPVW 360
 DB 301 LEAKTGKLGHOEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPVW 360
 QY 361 LNPEAGMWQCLLSDSGQVLESNIKVLPTWSTPVHPAPASALPAPPTGALPDPQTASALP 420
 DB 361 LNPEAGMWQCLLSDSGQVLESNIKVLPTWSTPVHPAPASALPAPPTGALPDPQTASALP 420
 QY 421 DPPASALPAPALAVISFLIGLGLGV-ACVLAATR 453
 DB 397 -----MALIVLGVAAGLLFLIGLGFPCVCRHR 425

RESULT 26
 US-08-466-368-2
 Sequence 2, Application US/08466368
 Patent No. 6093539

GENERAL INFORMATION:
 APPLICANT: Maddon, Paul J.
 APPLICANT: Littman, Dan R.
 APPLICANT: Chase, Leonard
 APPLICANT: Axel, Richard
 APPLICANT: Weiss, Robin
 APPLICANT: McDougal, J. S.
 TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN
 TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,368
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 24577-EI-B/JWP/ARC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-278-0400
 TELEFAX: 212-391-0525
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 394 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-466-368-2

Query Match 86.1%; Score 2001; DB 3; Length 394;
 Best Local Similarity 98.7%; Pred. No. 2.9e-159;
 Matches 389; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNRGVPPRHLLVQLALPPAATOGKVVYLGKGDVVELTCTASOKKSIOFHMKNNOIK 60
 DB 1 MNRGVPPRHLLVQLALPPAATOGKVVYLGKGDVVELTCTASOKKSIOFHMKNNOIK 60
 QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDYICEVEDQKEEVOL 120
 DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDYICEVEDQKEEVOL 120
 QY 121 LVFGLTANSPTHLLQGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
 DB 121 LVFGLTANSPTHLLQGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
 QY 181 TWCTVLQONQKVEFKIDIVLAFOKASSIYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
 DB 181 TWCTVLQONQKVEFKIDIVLAFOKASSIYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
 QY 241 QAEKASSKSWITFDLKNKEVSVKRVTQDPKLOMGKPLHLTLPOALPOYAGSGNLTLLA 300
 DB 241 QAEKASSKSWITFDLKNKEVSVKRVTQDPKLOMGKPLHLTLPOALPOYAGSGNLTLLA 300
 QY 301 LEAKTGKLGHOEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPVW 360
 DB 301 LEAKTGKLGHOEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPVW 360
 QY 361 LNPEAGMWQCLLSDSGQVLESNIKVLPTWSTPV 394
 DB 361 LNPEAGMWQCLLSDSGQVLESNIKVLPTWSTPV 394

RESULT 27
 US-08-328-500-2
 Sequence 2, Application US/08328500
 Patent No. 6673896

GENERAL INFORMATION:
 APPLICANT: Maddon, Paul J.
 APPLICANT: Sweet, Richard W.
 APPLICANT: Athos, James
 TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

```
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/328,500
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/24577-CY
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-328-500-2

Query Match      86.1%; Score 2001; DB 4: Length 394;
Best Local Similarity 98.7%; Pred. No. 2,9e-159;
Matches 389; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NMRGVFRLHLVLQALLPATQGNKVLGKKGDVVELTCTASQKSIQFHMKNNOIK 60
DB 1 NMRGVFRLHLVLQALLPATQGNKVLGKKGDVVELTCTASQKSIQFHMKNNOIK 60

QY 61 ILGNQGSFLTKGPPSKINDRADSRSLMDQGNFPLIINKIKIEDSPYICEVEDQKEEVL 120
DB 61 ILGNQGSFLTKGPPSKINDRADSRSLMDQGNFPLIINKIKIEDSPYICEVEDQKEEVL 120

QY 121 LVFGLTANSDTHLLOQSLTTLSPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLOQSLTTLSPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY 181 TWCTVLQNKQKVEFKIDIVLAFQKASSIYVKEGQVEFSPFLAFVTEKLTGSGELMW 240
DB 181 TWCTVLQNKQKVEFKIDIVLAFQKASSIYVKEGQVEFSPFLAFVTEKLTGSGELMW 240

QY 241 QAEKSSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGLTLA 300
DB 241 QAEKSSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGLTLA 300

QY 301 LEAKTGKLGHOEVLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKYSKEKPYVW 360
DB 301 LEAKTGKLGHOEVLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKYSKEKPYVW 360

QY 361 LNPEAGMMOCLLSDSGQVLLSNIKVLPTWSTPV 394
DB 361 LNPEAGMMOCLLSDSGQVLLSNIKVLPTWSTPV 394

RESULT 28
5223418-2
PATENT No. 5223418
APPLICANT: ARCURI, EDWARD J.;BRAUNER, MARY E.;DONOVAN, MARY
J.;GERBER, ROBERT G.;KELLER, JOHN A.
TITLE OF INVENTION: METHOD OF IMPROVING THE YIELD OF
HETEROLOGOUS PROTEINS PRODUCED BY STREPTOMYCES LIVIDANS
NUMBER OF SEQUENCES: 2
CURRENT APPLICATION DATA:
REGISTRATION NUMBER: US/07/589,979
FILING DATE: 28-SEP-1990
SEQ ID NO:2
LENGTH: 394
5223418-2

Query Match      83.9%; Score 1951; DB 6: Length 394;
Best Local Similarity 97.5%; Pred. No. 4,4e-155;
Matches 384; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 NMRGVFRLHLVLQALLPATQGNKVLGKKGDVVELTCTASQKSIQFHMKNNOIK 60
```

```
DB 1 NMRGVFRLHLVLQALLPATQGNKVLGKKGDVVELTCTASQKSIQFHMKNNOIK 60

QY 61 ILGNQGSFLTKGPPSKINDRADSRSLMDQGNFPLIINKIKIEDSPYICEVEDQKEEVL 120
DB 61 ILGNQGSFLTKGPPSKINDRADSRSLMDQGNFPLIINKIKIEDSPYICEVEDQKEEVL 120

QY 121 LVFGLTANSDTHLLOQSLTTLSPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLOQSLTTLSPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY 181 TWCTVLQNKQKVEFKIDIVLAFQKASSIYVKEGQVEFSPFLAFVTEKLTGSGELMW 240
DB 181 TWCTVLQNKQKVEFKIDIVLAFQKASSIYVKEGQVEFSPFLAFVTEKLTGSGELMW 240

QY 241 QAEKSSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGLTLA 300
DB 241 QAEKSSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGLTLA 300

QY 301 LEAKTGKLGHOEVLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKYSKEKPYVW 360
DB 301 LEAKTGKLGHOEVLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKYSKEKPYVW 360

QY 361 LNPEAGMMOCLLSDSGQVLLSNIKVLPTWSTPV 394
DB 361 LNPEAGMMOCLLSDSGQVLLSNIKVLPTWSTPV 394

RESULT 29
US-08-867-149-1
Sequence 1, Application US/08867149
Patent No. 5912176
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Antibody Against a Host Cell
TITLE OF INVENTION: Antigen Complex for Pre- and Post-Exposure
PROTECTION FROM INFECTION BY HIV PRIMARY ISOLATES
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,149
FILING DATE: 28-Feb-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/657,149
FILING DATE: 03-June-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4145
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-867-149-1
```

Query Match 82.1%; Score 1909; DB 2; Length 433;
Best Local Similarity 88.6%; Pred. No. 1.6e-15;
Matches 380; Conservative 2; Mismatches 17; Indels 30; Gaps 2;

QY 26 NKVVLGKGGDTVELTCTASQKSIQPHWKNNOIKILGNQGSFLTGPSKLNDRADSRRS 85
DB 1 NKVVLGKGGDTVELTCTASQKSIQPHWKNNOIKILGNQGSFLTGPSKLNDRADSRRS 60
QY 86 LMDQNPFLITKNIKIEDSDTYICEVEDQKEEVLVFGLTANSPTHLLQGSLTLTLES 145
DB 61 LMDQNPFLITKNIKIEDSDTYICEVEDQKEEVLVFGLTANSPTHLLQGSLTLTLES 120
QY 146 PGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLOKQKVEFKIDIVLAFQ 205
DB 121 PGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLOKQKVEFKIDIVLAFQ 180
QY 206 KASSIVYKKEGEQVEFSFPLAFVTEKLTGSGELMWQAEKASSKSWITFDLKNKEVSVKR 265
DB 181 KASSIVYKKEGEQVEFSFPLAFVTEKLTGSGELMWQAEKASSKSWITFDLKNKEVSVKR 240
QY 266 VTQDPKLOMGKKLPPLHLTPQALPOYAGSGNLTALAEKTKGLHOEVNLVVMRATOLQKN 325
DB 241 VTQDPKLOMGKKLPPLHLTPQALPOYAGSGNLTALAEKTKGLHOEVNLVVMRATOLQKN 300
QY 326 LTCEVWGPTSPKMLSLKLENKEAKVSKREKPVWVLPNPEAGMOCCLISDSGOVLLESNIK 385
DB 301 LTCEVWGPTSPKMLSLKLENKEAKVSKREKPVWVLPNPEAGMOCCLISDSGOVLLESNIK 360
QY 386 VLPTWSTVPHRASALPAPPTGSALPDPTQASALPDPPASALPALAVISFLGLGLGV 445
DB 361 VLPTWSTVPHRASALPAPPTGSALPDPTQASALPDPPASALPALAVISFLGLGLGV 391
QY 446 -ACVLARTR 453
DB 392 FFCVRCRHR 400

RESULT 30

US-08-808-374-1
Sequence 1, Application US/08808374
Patent No. 5961976

GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Antibody Against a Host Cell
TITLE OF INVENTION: Antigen Complex for Pre- and Post-Exposure
TITLE OF INVENTION: Protection from Infection by HIV Primary Isolates
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,374
FILING DATE: 28-Feb-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/657,149
FILING DATE: 03-June-1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4145

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)415-8745
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-808-374-1

Query Match 82.1%; Score 1909; DB 2; Length 433;
Best Local Similarity 88.6%; Pred. No. 1.6e-15;
Matches 380; Conservative 2; Mismatches 17; Indels 30; Gaps 2;

QY 26 NKVVLGKGGDTVELTCTASQKSIQPHWKNNOIKILGNQGSFLTGPSKLNDRADSRRS 85
DB 1 NKVVLGKGGDTVELTCTASQKSIQPHWKNNOIKILGNQGSFLTGPSKLNDRADSRRS 60
QY 86 LMDQNPFLITKNIKIEDSDTYICEVEDQKEEVLVFGLTANSPTHLLQGSLTLTLES 145
DB 61 LMDQNPFLITKNIKIEDSDTYICEVEDQKEEVLVFGLTANSPTHLLQGSLTLTLES 120
QY 146 PGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLOKQKVEFKIDIVLAFQ 205
DB 121 PGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLOKQKVEFKIDIVLAFQ 180
QY 206 KASSIVYKKEGEQVEFSFPLAFVTEKLTGSGELMWQAEKASSKSWITFDLKNKEVSVKR 265
DB 181 KASSIVYKKEGEQVEFSFPLAFVTEKLTGSGELMWQAEKASSKSWITFDLKNKEVSVKR 240
QY 266 VTQDPKLOMGKKLPPLHLTPQALPOYAGSGNLTALAEKTKGLHOEVNLVVMRATOLQKN 325
DB 241 VTQDPKLOMGKKLPPLHLTPQALPOYAGSGNLTALAEKTKGLHOEVNLVVMRATOLQKN 300
QY 326 LTCEVWGPTSPKMLSLKLENKEAKVSKREKPVWVLPNPEAGMOCCLISDSGOVLLESNIK 385
DB 301 LTCEVWGPTSPKMLSLKLENKEAKVSKREKPVWVLPNPEAGMOCCLISDSGOVLLESNIK 360
QY 386 VLPTWSTVPHRASALPAPPTGSALPDPTQASALPDPPASALPALAVISFLGLGLGV 445
DB 361 VLPTWSTVPHRASALPAPPTGSALPDPTQASALPDPPASALPALAVISFLGLGLGV 391
QY 446 -ACVLARTR 453
DB 392 FFCVRCRHR 400

RESULT 31

US-09-100-409A-1
Sequence 1, Application US/09100409A
Patent No. 6090388

GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
TITLE OF INVENTION: IMMUNE DISORDERS
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A

```

; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-409A-1

Query Match      81.9%; Score 1904; DB 3; Length 433;
Best Local Similarity 88.6%; Pred. No. 4.3e-151;
Matches 379; Conservative 2; Mismatches 17; Indels 30; Gaps 2;

QY 27 KVLGKGGDTVELTCTASQKSIQPHWKNKSNQIKLGNQSGFLTKGPSKLNDRADSRSL 86
DB 2 KVLGKGGDTVELTCTASQKSIQPHWKNKSNQIKLGNQSGFLTKGPSKLNDRADSRSL 61
QY 87 WQGNPPLIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLLQGSLLTLTLESP 146
DB 62 WQGNPPLIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLLQGSLLTLTLESP 121
QY 147 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLOKQKVEFKIDIVVLAFOK 206
DB 122 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLOKQKVEFKIDIVVLAFOK 181
QY 207 ASSIYKKEGGEVSFPPLAFTVEKLTGSGELMWAERASSKSWITFDLNKKEVSVKRV 266
DB 182 ASSIYKKEGGEVSFPPLAFTVEKLTGSGELMWAERASSKSWITFDLNKKEVSVKRV 241
QY 267 TODPPLQMGKKLPLHLTLPQALPOYAGSGNLTALAEATGKLGHOENVLVMMRATOLQ 326
DB 242 TODPPLQMGKKLPLHLTLPQALPOYAGSGNLTALAEATGKLGHOENVLVMMRATOLQ 301
QY 327 TCEVWGPTSPKMLSLKLENKEAKYSKREKPYWVLNPEAGMOCCLSDSGVLLSESNIKV 386
DB 302 TCEVWGPTSPKMLSLKLENKEAKYSKREKPYWVLNPEAGMOCCLSDSGVLLSESNIKV 361
QY 387 LFTWSTPVHPRASALPAPPTGSALPDPTQASALPDPPAASALPALAVISFLGLGLGV- 445
DB 362 LFTWSTPVQP-----MALIVGVAAGLLFLGLGIF 392
QY 446 ACVLARTR 453
DB 393 FCVRCRHR 400

RESULT 32
US-08-236-311-4
; Sequence 4, Application US/08236311
; Patent No. 5565335
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; APPLICANT: Gregory, Timothy J.
; TITLE OF INVENTION: Adhesion Variants
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/236,311
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/936190
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/842777
; FILING DATE: 18-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/250785
; FILING DATE: 28-SEP-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 444PIC2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-236-311-4

Query Match      81.9%; Score 1904; DB 1; Length 434;
Best Local Similarity 99.7%; Pred. No. 4.3e-151;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 24 QGNKVVLGKKDDTVELTCTASQKSIQPHWKNKSNQIKLGNQSGFLTKGPSKLNDRASR 83
DB 56 QGNKVVLGKKDDTVELTCTASQKSIQPHWKNKSNQIKLGNQSGFLTKGPSKLNDRASR 115
QY 84 RSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLLQGSLLTLT 143
DB 116 RSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLLQGSLLTLT 175
QY 144 ESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLOKQKVEFKIDIVVLA 203
DB 176 ESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLOKQKVEFKIDIVVLA 235
QY 204 FOKASSIYKKEGGEVSFPPLAFTVEKLTGSGELMWAERASSKSWITFDLNKKEVS 263
DB 236 FOKASSIYKKEGGEVSFPPLAFTVEKLTGSGELMWAERASSKSWITFDLNKKEVS 295
QY 264 KRVTDPLQMGKKLPLHLTLPQALPOYAGSGNLTALAEATGKLGHOENVLVMMRATOLQ 323
DB 296 KRVTDPLQMGKKLPLHLTLPQALPOYAGSGNLTALAEATGKLGHOENVLVMMRATOLQ 355
QY 324 KNLTCVWGPTSPKMLSLKLENKEAKYSKREKPYWVLNPEAGMOCCLSDSGVLLSESN 383
DB 356 KNLTCVWGPTSPKMLSLKLENKEAKYSKREKPYWVLNPEAGMOCCLSDSGVLLSESN 415
QY 384 IKVLPWTSTP 393
DB 416 IKVLPWTSTP 425

RESULT 33
US-08-457-918-4
; Sequence 4, Application US/08457918
; Patent No. 6117655
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; APPLICANT: Gregory, Timothy J.

```

```

/ TITLE OF INVENTION: Adhesion Variants
/ NUMBER OF SEQUENCES: 25
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: Genentech, Inc.
/ STREET: 460 Point San Bruno Blvd
/ CITY: South San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94080
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: patin (Genentech)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/457,918
/ FILING DATE: 1-JUN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/236311
/ FILING DATE: 02-MAY-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/936190
/ FILING DATE: 26-AUG-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/842777
/ FILING DATE: 18-FEB-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/250785
/ FILING DATE: 28-SEP-1988
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/104329
/ FILING DATE: 02-OCT-1987
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kubinec, Jeffrey S.
/ REGISTRATION NUMBER: 36,575
/ REFERENCE/DOCKET NUMBER: P0444PIC3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415/225-8228
/ TELEFAX: 415/952-9881
/ TELEX: 910371-7168
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 434 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ US-08-457-918-4

Query Match      81.9%; Score 1904; DB 3; Length 434;
Best Local Similarity 99.7%; Pred. No. 4.3e-151;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 24 QGNVYVGGKGDVVELTCTASQKSIQPHMKNNOIKILGNQGSFLTGPBKLNDRADSR 83
Db 56 QGNVYVGGKGDVVELTCTASQKSIQPHMKNNOIKILGNQGSFLTGPBKLNDRADSR 115
Oy 84 RSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDFTHLQGSILTLT 143
Db 116 RSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDFTHLQGSILTLT 175
Oy 144 ESPGSSPSVOCRSPPRGKNIQGGKTLVSQLELDOSGTWCTVLOKQKVEFKIDIVVLA 203
Db 176 ESPGSSPSVOCRSPPRGKNIQGGKTLVSQLELDOSGTWCTVLOKQKVEFKIDIVVLA 235
Oy 204 FOKASSIYKKEGEVEFSFPLAFVEKLTGSGELMWQAEBASSKSWITFDLNKKEVS 263
Db 236 FOKASSIYKKEGEVEFSFPLAFVEKLTGSGELMWQAEBASSKSWITFDLNKKEVS 295
Oy 264 KRVTQDPKIQMGKPLHLTLPOALPOYAGSGNLTALAEKTKLHQBENLVYMRATQLO 323
Db 296 KRVTQDPKIQMGKPLHLTLPOALPOYAGSGNLTALAEKTKLHQBENLVYMRATQLO 355
Oy 324 KNLTCCEVMGPTSPKLMLSLKLENKEAKVSKREKPYVVLNBPAGMOCCLSDSGVLL 383
```

```

Db 356 KNLTCCEVMGPTSPKLMLSLKLENKEAKVSKREKPYVVLNBPAGMOCCLSDSGVLL 415
Oy 384 IKVLPMTWSTP 393
Db 416 IKVLPMTWSTP 425

RESULT 34
5171838-13
/ Patent No. 5171838
/ APPLICANT: CHIBA, YUKINOBU
/ TITLE OF INVENTION: LEU3A BINDING PEPTIDES
/ NUMBER OF SEQUENCES: 24
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/526,921
/ FILING DATE: 22-MAY-1990
/ SEQ ID NO:13:
/ LENGTH: 433
5171838-13
```

Query Match 73.6%; Score 1712; DB 6; Length 433;

Best Local Similarity 81.6%; Pred. No. 5e-135; Matches 355; Conservative 8; Mismatches 28; Indels 44; Gaps 6;

```

Oy 27 KVVLGKKGDVVELTCTASQKSIQPHMKNNOIKILGNQGSFLTGPBKLNDRADSR 85
Db 2 KVVLGKKGDVVELTCTASQKSIQPHMKNNOIKILGNQGSFLTGPBKLNDRADSR 61
Oy 86 ----LMDQGNFPLIINKLKIED-SDTYICEVEDQKEEVQLVFGLTANSDFTHLQGS 139
Db 62 NQIKILGNKGSF-LTGKPSKLNDRADS-----RSEVQLVFGLTANSDFTHLQGS 114
Oy 140 TLTLSPGSSPSVOCRSPPRGKNIQGGKTLVSQLELDOSGTWCTVLOKQKVEFKID 199
Db 115 TLTLSPGSSPSVOCRSPPRGKNIQGGKTLVSQLELDOSGTWCTVLOKQKVEFKID 174
Oy 200 VVLAFOKASSIYKKEGEVEFSFPLAFVEKLTGSGELMWQAEBASSKSWITFDLNK 259
Db 175 VVLAFOKASSIYKKEGEVEFSFPLAFVEKLTGSGELMWQAEBASSKSWITFDLNK 234
Oy 260 EVSVKRYTQDPKIQMGKPLHLTLPOALPOYAGSGNLTALAEKTKLHQBENLVY 319
Db 235 EVSVKRYTQDPKIQMGKPLHLTLPOALPOYAGSGNLTALAEKTKLHQBENLVY 294
Oy 320 TOLQKNLTCEVMGPTSPKLMLSLKLENKEAKVSKREKPYVVLNBPAGMOCCLSD 379
Db 295 TOLQKNLTCEVMGPTSPKLMLSLKLENKEAKVSKREKPYVVLNBPAGMOCCLSD 354
Oy 380 LESNIKVLPTWSTPVHPRASALPAPPGSALPDPTASALPDPPAASALPALAVIS 439
Db 355 LESNIKVLPTWSTPVHPRASALPAPPGSALPDPTASALPDPPAASALPALAVIS 385
Oy 440 GLGIGV-ACVLAATR 453
Db 386 FIGIGIFCVCRCHR 400
```

```

RESULT 35
5223394-11
/ Patent No. 5223394
/ APPLICANT: WALLNER, BARBARA
/ TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
/ LYMPHOCTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
/ LINKAGE SIGNAL SEQUENCE
/ NUMBER OF SEQUENCES: 12
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/335,688
/ FILING DATE: 10-APR-1989
/ SEQ ID NO:11:
/ LENGTH: 318
5223394-11
```

```
Query Match          58.8%; Score 1363; DB 6; Length 318;
Best Local Similarity 93.4%; Pred. No. 1.9e-106;
Matches 268; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY      1 MNRGVPFRHLILVQLALIPATOGKVVLAGKGGDTVELTCTASQKSIQFHMKNNOIK 60
        |||
DB      1 MNRGVPFRHLILVQLALIPATOGKVVLAGKGGDTVELTCTASQKSIQFHMKNNOIK 60

QY      61 ILNGGSPFTKGPSTKLNDRADSRSLMDGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
        |||
DB      61 ILNGGSPFTKGPSTKLNDRADSRSLMDGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120

QY      121 LVFGILTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELODSG 180
        |||
DB      121 LVFGILTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELODSG 180

QY      181 TWTCITVLONQKKEVEFKIDIVLAFQKASSIVYKKEGEQVEFSFPLAFVTEKLTSGGELMW 240
        |||
DB      181 TWTCITVLONQKKEVEFKIDIVLAFQKASSIVYKKEGEQVEFSFPLAFVTEKLTSGGELMW 240

QY      241 QAEKASSSSKSWITTFDLKNKEVSVKRVTOQPKLOMGKPLHLTLPOA 287
        |||
DB      241 QAEKASSSSKSWITTFDLKNKEVSVKRVISNPLNNTTSSILTTCTIPSS 287

RESULT 36
5223394-9
; Patent No. 5223394
; APPLICANT: WALINER, BARBARA
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
; LINKAGE SIGNAL SEQUENCE
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/335,688
; FILING DATE: 10-APR-1989
; SEQ ID NO:9:
; LENGTH: 295
5223394-9

Query Match          58.6%; Score 1363; DB 6; Length 295;
Best Local Similarity 99.3%; Pred. No. 4.5e-106;
Matches 266; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 MNRGVPFRHLILVQLALIPATOGKVVLAGKGGDTVELTCTASQKSIQFHMKNNOIK 60
        |||
DB      1 MNRGVPFRHLILVQLALIPATOGKVVLAGKGGDTVELTCTASQKSIQFHMKNNOIK 60

QY      61 ILNGGSPFTKGPSTKLNDRADSRSLMDGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
        |||
DB      61 ILNGGSPFTKGPSTKLNDRADSRSLMDGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120

QY      121 LVFGILTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELODSG 180
        |||
DB      121 LVFGILTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELODSG 180

QY      181 TWTCITVLONQKKEVEFKIDIVLAFQKASSIVYKKEGEQVEFSFPLAFVTEKLTSGGELMW 240
        |||
DB      181 TWTCITVLONQKKEVEFKIDIVLAFQKASSIVYKKEGEQVEFSFPLAFVTEKLTSGGELMW 240

QY      241 QAEKASSSSKSWITTFDLKNKEVSVKRVTOQPKLOMGKPLHLTLPOA 287
        |||
DB      241 QAEKASSSSKSWITTFDLKNKEVSVKRVTR 268

RESULT 37
US-08-477-4608-6
; Sequence 6, Application US/084774608
; Patent No. 6034223
; GENERAL INFORMATION:
; APPLICANT: Progenice Pharmaceuticals, Inc.
; TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
; TITLE OF INVENTION: CD4-GAMMA2 AND CD4-1GG2 IMMUNOCONJUGATES, AND USES THEREOF
```

```
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,4608
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte

US-08-477-4608-6

Query Match          45.2%; Score 1050.5; DB 3; Length 310;
Best Local Similarity 75.4%; Pred. No. 6.8e-80;
Matches 221; Conservative 12; Mismatches 41; Indels 19; Gaps 4;

QY      1 MNRGVPFRHLILVQLALIPATOGKVVLAGKGGDTVELTCTASQKSIQFHMKNNOIK 60
        |||
DB      1 MNRGVPFRHLILVQLALIPATOGKVVLAGKGGDTVELTCTASQKSIQFHMKNNOIK 60

QY      61 ILNGGSPFTKGPSTKLNDRADSRSLMDGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
        |||
DB      61 ILNGGSPFTKGPSTKLNDRADSRSLMDGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120

QY      121 LVFGILTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELODSG 180
        |||
DB      121 LVFGILTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELODSG 180

QY      181 TWTCITVLONQKKEVEFKIDIVLAFQKASSIVYKKEGEQVEFSFPLAFVTEKLTSGGELMW 240
        |||
DB      181 TWTCITVLONQKKEVEFKIDIVLAFQKASSIVYKKEGEQVEFSFPLAFVTEKLTSGGELMW 240

QY      241 QAEKASSSSKSWITTFDLKNKEVSVKRVTOQPKLOMGKPLHLTLPOA 287
        |||
DB      241 QAEKASSSSKSWITTFDLKNKEVSVKRVTOQPKLOMGKPLHLTLPOA 287

RESULT 38
US-08-379-516-6
; Sequence 6, Application US/08379516
; Patent No. 6083478
; GENERAL INFORMATION:
; APPLICANT: Allaway, Graham P.
; TITLE OF INVENTION: Moietly-Conjugated CD4-Gamma2 and CD4-1GG2
```

TITLE OF INVENTION: Immunocjugates and Uses thereof
FILE REFERENCE: 41215-A-PCT-US
CURRENT APPLICATION NUMBER: US/08/379,516
CURRENT FILING DATE: 1996-06-10
EARLIER APPLICATION NUMBER: PCT/US93/07422
EARLIER FILING DATE: 1993-08-06
EARLIER APPLICATION NUMBER: 07/927,931
EARLIER FILING DATE: 1992-08-07
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 6
LENGTH: 310
TYPE: PRT
ORGANISM: Homo sapiens
US-08-379-516-6

Query Match 45.2%; Score 1050.5; DB 3; Length 310;
Best Local Similarity 75.4%; Pred. No. 6.8e-80;
Matches 221; Conservative 12; Mismatches 41; Indels 19; Gaps 4;

Qy 1 MNRGVPFRHLVLTQALLPATQGNKVVLGKGGDTVELTCTASOKKSIQPHMKNQIK 60
Db 1 MNRGVPFRHLVLTQALLPATQGNKVVLGKGGDTVELTCTASOKKSIQPHMKNQIK 60
Qy 61 ILNGGSFLLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQL 120
Db 61 ILNGGSFLLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQL 120
Qy 121 LVFGTANSDPHLLQGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGTANSDPHLLQGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy 181 TWTCVTQONQKVEFKIDIVLAFQKA--SSIVYKKEGEQVEFPLAFVETKLTG---- 234
Db 181 TWTCVTQONQKVEFKIDIVLAFQKA--SSIVYKKEGEQVEFPLAFVETKLTG---- 234
Qy 235 SGEIWMQAEASSSKSWITFDLKNKEVSVKRVTPDPLQMGKPLPLHLTPQA 287
Db 235 SGEIWMQAEASSSKSWITFDLKNKEVSVKRVTPDPLQMGKPLPLHLTPQA 287
Qy 239 EAKYQWKVDNALQSGN-----SQESVTEQDSKDSSTYSLSSTLTLSKA 280
Db 239 EAKYQWKVDNALQSGN-----SQESVTEQDSKDSSTYSLSSTLTLSKA 280

RESULT 39
US-09-329-916-6
Sequence 6, Application US/09329916
Patent No. 6177549
GENERAL INFORMATION:
APPLICANT: Progenics Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/329,916
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/477,460
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.

REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-09-329-916-6

Query Match 45.2%; Score 1050.5; DB 3; Length 310;
Best Local Similarity 75.4%; Pred. No. 6.8e-80;
Matches 221; Conservative 12; Mismatches 41; Indels 19; Gaps 4;

Qy 1 MNRGVPFRHLVLTQALLPATQGNKVVLGKGGDTVELTCTASOKKSIQPHMKNQIK 60
Db 1 MNRGVPFRHLVLTQALLPATQGNKVVLGKGGDTVELTCTASOKKSIQPHMKNQIK 60
Qy 61 ILNGGSFLLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQL 120
Db 61 ILNGGSFLLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQL 120
Qy 121 LVFGTANSDPHLLQGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGTANSDPHLLQGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy 181 TWTCVTQONQKVEFKIDIVLAFQKA--SSIVYKKEGEQVEFPLAFVETKLTG---- 234
Db 181 TWTCVTQONQKVEFKIDIVLAFQKA--SSIVYKKEGEQVEFPLAFVETKLTG---- 234
Qy 235 SGEIWMQAEASSSKSWITFDLKNKEVSVKRVTPDPLQMGKPLPLHLTPQA 287
Db 235 SGEIWMQAEASSSKSWITFDLKNKEVSVKRVTPDPLQMGKPLPLHLTPQA 287
Qy 239 EAKYQWKVDNALQSGN-----SQESVTEQDSKDSSTYSLSSTLTLSKA 280
Db 239 EAKYQWKVDNALQSGN-----SQESVTEQDSKDSSTYSLSSTLTLSKA 280

RESULT 40
US-08-485-372A-6
Sequence 6, Application US/08485372A
Patent No. 6187748
GENERAL INFORMATION:
APPLICANT: Beaudry, Gary A.
TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IGG2 IMMUNOCJUGATES, AND USES THEREOF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,372A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,227
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.

REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 37690-II-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-08-485-372A-6

Query Match 45.2%; Score 1050.5; DB 3; Length 310;
Best Local Similarity 75.4%; Pred. No. 6.8e-80;
Matches 221; Conservative 12; Mismatches 41; Indels 19; Gaps 4;

QY 1 MNRGVPRHLLVQLALPAATQGNKVVLGKGGDTVELTCTASQKSIQFHMKNNOIK 60
DB 1 MNRGVPRHLLVQLALPAATQGNKVVLGKGGDTVELTCTASQKSIQFHMKNNOIK 60
QY 61 ILNGGSLFTGKPSKLNDRADSRSLMDQGNFPLIKLKIEDSTYICEVEDQKEEYOL 120
DB 61 ILNGGSLFTGKPSKLNDRADSRSLMDQGNFPLIKLKIEDSTYICEVEDQKEEYOL 120
QY 121 LVFGTANSDTHLQOGSLTTLSPGSSPVOCRSPRGKNIQGGKTLVSQLELODSG 180
DB 121 LVFGTANSDTHLQOGSLTTLSPGSSPVOCRSPRGKNIQGGKTLVSQLELODSG 180
QY 181 TWTCVTLQNKQKVEFKIDIVLAFOKA--SSIVYKKEGEQVEFSPLAFTVEKLTG--- 234
DB 181 TWTCVTLQNKQKVEFKIDIVLAFTVAAPSVFIFFPSDEQLKSG--TASVCLLNPFYR 238
QY 235 SGEIWMQERASSKSWITFDLKNKEVSKRYTQDPKIQMGKKLPLHLTLPOA 287
DB 235 EAKVQMKVDNALQSGN-----SQESVTEQDSKDYSLSTLTLSKA 280

RESULT 41
US-09-409-006A-6
Sequence 6, Application US/09409006A
Patent No. 6342586
GENERAL INFORMATION:
APPLICANT: Progenice Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/409,006A
FILING DATE: 29-SEP-1999
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.

REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/AMJ
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-09-409-006A-6

Query Match 45.2%; Score 1050.5; DB 4; Length 310;
Best Local Similarity 75.4%; Pred. No. 6.8e-80;
Matches 221; Conservative 12; Mismatches 41; Indels 19; Gaps 4;

QY 1 MNRGVPRHLLVQLALPAATQGNKVVLGKGGDTVELTCTASQKSIQFHMKNNOIK 60
DB 1 MNRGVPRHLLVQLALPAATQGNKVVLGKGGDTVELTCTASQKSIQFHMKNNOIK 60
QY 61 ILNGGSLFTGKPSKLNDRADSRSLMDQGNFPLIKLKIEDSTYICEVEDQKEEYOL 120
DB 61 ILNGGSLFTGKPSKLNDRADSRSLMDQGNFPLIKLKIEDSTYICEVEDQKEEYOL 120
QY 121 LVFGTANSDTHLQOGSLTTLSPGSSPVOCRSPRGKNIQGGKTLVSQLELODSG 180
DB 121 LVFGTANSDTHLQOGSLTTLSPGSSPVOCRSPRGKNIQGGKTLVSQLELODSG 180
QY 181 TWTCVTLQNKQKVEFKIDIVLAFOKA--SSIVYKKEGEQVEFSPLAFTVEKLTG--- 234
DB 181 TWTCVTLQNKQKVEFKIDIVLAFTVAAPSVFIFFPSDEQLKSG--TASVCLLNPFYR 238
QY 235 SGEIWMQERASSKSWITFDLKNKEVSKRYTQDPKIQMGKKLPLHLTLPOA 287
DB 235 EAKVQMKVDNALQSGN-----SQESVTEQDSKDYSLSTLTLSKA 280

RESULT 42
US-08-484-681-6
Sequence 6, Application US/08484681
Patent No. 6451313
GENERAL INFORMATION:
APPLICANT: Beaudry, Gary A.
TITLE OF INVENTION: CD4-GAMMA2 CD4-IGG2 CHIMERAS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,681
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 37690-II-B
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-08-484-681-6

Query Match 45.2%; Score 1050.5; DB 4; Length 310;
Best Local Similarity 75.4%; Pred. No. 6.8e-80;
Matches 221; Conservative 12; Mismatches 41; Indels 19; Gaps 4;

QY 1 MNRGVPFRHLVLVQLALPPATQGNKVVLGKGDVVELTCTASOKKSIQFHWKNSNQIK 60
DB 1 MNRGVPFRHLVLVQLALPPATQGNKVVLGKGDVVELTCTASOKKSIQFHWKNSNQIK 60
QY 61 ILGNQGSFLTKGSKLNDRADRSRLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSKLNDRADRSRLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELQDSG 180
DB 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELQDSG 180
QY 181 TWTCTVLQNGKVEFKIDIVLAFQKA--SSIVYKKEGEQVEFSPLAFTVEKLTG---- 234
DB 181 TWTCTVLQNGKVEFKIDIVLAFQKA--SSIVYKKEGEQVEFSPLAFTVEKLTG---- 234
QY 235 SGELMWQAEKRASSKSMITFDLKNKEVSVKRYTQDPLQMGKGLPLHLTLPQA 287
DB 235 SGELMWQAEKRASSKSMITFDLKNKEVSVKRYTQDPLQMGKGLPLHLTLPQA 287
QY 239 EAKVQWKVDNALQSGN-----SQESVTEQDSKDSYLSSTLTLSKA 280
DB 239 EAKVQWKVDNALQSGN-----SQESVTEQDSKDSYLSSTLTLSKA 280

RESULT 43

PCT-US93-07422-6
Sequence 6, Application PC/TUS9307422
GENERAL INFORMATION:
APPLICANT: Progenics Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IGG2 IMMUNOCONJUGATES, AND USES THEREOF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07422
FILING DATE: 19930806
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550

TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
PCT-US93-07422-6

Query Match 45.2%; Score 1050.5; DB 5; Length 310;
Best Local Similarity 75.4%; Pred. No. 6.8e-80;
Matches 221; Conservative 12; Mismatches 41; Indels 19; Gaps 4;

QY 1 MNRGVPFRHLVLVQLALPPATQGNKVVLGKGDVVELTCTASOKKSIQFHWKNSNQIK 60
DB 1 MNRGVPFRHLVLVQLALPPATQGNKVVLGKGDVVELTCTASOKKSIQFHWKNSNQIK 60
QY 61 ILGNQGSFLTKGSKLNDRADRSRLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSKLNDRADRSRLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELQDSG 180
DB 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELQDSG 180
QY 181 TWTCTVLQNGKVEFKIDIVLAFQKA--SSIVYKKEGEQVEFSPLAFTVEKLTG---- 234
DB 181 TWTCTVLQNGKVEFKIDIVLAFQKA--SSIVYKKEGEQVEFSPLAFTVEKLTG---- 234
QY 235 SGELMWQAEKRASSKSMITFDLKNKEVSVKRYTQDPLQMGKGLPLHLTLPQA 287
DB 235 SGELMWQAEKRASSKSMITFDLKNKEVSVKRYTQDPLQMGKGLPLHLTLPQA 287
QY 239 EAKVQWKVDNALQSGN-----SQESVTEQDSKDSYLSSTLTLSKA 280
DB 239 EAKVQWKVDNALQSGN-----SQESVTEQDSKDSYLSSTLTLSKA 280

RESULT 44

US-08-477-460B-4
Sequence 4, Application US/08477460B
Patent No. 6034223
GENERAL INFORMATION:
APPLICANT: Progenics Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IGG2 IMMUNOCONJUGATES, AND USES THEREOF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,460B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550

TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-08-477-460B-4

Query Match 45.0%; Score 1046; DB 3; Length 530;
Best Local Similarity 93.4%; Pred. No. 3.6e-79;
Matches 211; Conservative 2; Mismatches 3; Indels 10; Gaps 3;

QY 1 MNRGVPFRHLVLVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQPHMKNNOIK 60
DB 1 MNRGVPFRHLVLVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQPHMKNNOIK 60
QY 61 ILNGSGSLTGGPSKLNDRADSRSLMDQGNFPLIIKULKIEDSDTYICEVEDQKEVQL 120
DB 61 ILNGSGSLTGGPSKLNDRADSRSLMDQGNFPLIIKULKIEDSDTYICEVEDQKEVQL 120
QY 121 LVFGILTANSDTHLLOQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
DB 121 LVFGILTANSDTHLLOQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
QY 181 TWTCTVLONQKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLA 226
DB 181 TWTCTVLONQKVEFKIDIVLAF--AST-----KGPSTV---FPLA 216

RESULT 45
US-08-379-516-4

Sequence 4, Application US/08379516
Patent No. 5083478
GENERAL INFORMATION:
APPLICANT: Allaway, Graham P.
REFERENCE: Maddon, Paul J.
TITLE OF INVENTION: No. 6083478-peptidyl moiety-conjugated CD4-Gamma2 and CD4-1G62
TITLE OF INVENTION: Immunocjugates and Uses Thereof
FILE REFERENCE: 41215-A-PCT-US
CURRENT APPLICATION NUMBER: US/08/379,516
CURRENT FILING DATE: 1996-06-10
EARLIER APPLICATION NUMBER: PCT/US93/07422
EARLIER FILING DATE: 1993-08-06
EARLIER APPLICATION NUMBER: 07/927,931
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 4
LENGTH: 530
TYPE: PRT
ORGANISM: Homo sapiens
US-08-379-516-4

Query Match 45.0%; Score 1046; DB 3; Length 530;
Best Local Similarity 93.4%; Pred. No. 3.6e-79;
Matches 211; Conservative 2; Mismatches 3; Indels 10; Gaps 3;

QY 1 MNRGVPFRHLVLVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQPHMKNNOIK 60
DB 1 MNRGVPFRHLVLVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQPHMKNNOIK 60
QY 61 ILNGSGSLTGGPSKLNDRADSRSLMDQGNFPLIIKULKIEDSDTYICEVEDQKEVQL 120
DB 61 ILNGSGSLTGGPSKLNDRADSRSLMDQGNFPLIIKULKIEDSDTYICEVEDQKEVQL 120
QY 121 LVFGILTANSDTHLLOQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
DB 121 LVFGILTANSDTHLLOQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180

DB 121 LVFGILTANSDTHLLOQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
QY 181 TWTCTVLONQKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLA 226
DB 181 TWTCTVLONQKVEFKIDIVLAF--AST-----KGPSTV---FPLA 216

RESULT 46
US-09-329-916-4

Sequence 4, Application US/09329916
Patent No. 617549
GENERAL INFORMATION:
APPLICANT: Progenics Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/329,916
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/477,460
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-09-329-916-4

Query Match 45.0%; Score 1046; DB 3; Length 530;
Best Local Similarity 93.4%; Pred. No. 3.6e-79;
Matches 211; Conservative 2; Mismatches 3; Indels 10; Gaps 3;

QY 1 MNRGVPFRHLVLVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQPHMKNNOIK 60
DB 1 MNRGVPFRHLVLVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQPHMKNNOIK 60
QY 61 ILNGSGSLTGGPSKLNDRADSRSLMDQGNFPLIIKULKIEDSDTYICEVEDQKEVQL 120
DB 61 ILNGSGSLTGGPSKLNDRADSRSLMDQGNFPLIIKULKIEDSDTYICEVEDQKEVQL 120
QY 121 LVFGILTANSDTHLLOQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
DB 121 LVFGILTANSDTHLLOQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180

Qy 181 TWCTCTVLQNGKQKVEFKIDIVLAFQKASSIYKKEGEVSFPPLA 226
Db 181 TWCTCTVLQNGKQKVEFKIDIVLAF--AST-----KGPSV---FPLA 216

RESULT 47
US-08-485-372A-4
Sequence 4, Application US/08485372A
Patent No. 6187748
GENERAL INFORMATION:
APPLICANT: Beaudry, Gary A.
APPLICANT: Maddon, Paul J.
TITLE OF INVENTION: CD4-GAMMA2 CD4-IGG2 CHIMERAS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,372A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,227
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 37690-II-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-08-485-372A-4

Query Match 45.0%; Score 1046; DB 3; Length 530;
Best Local Similarity 93.4%; Pred. No. 3,6e-79;
Matches 211; Conservative 2; Mismatches 3; Indels 10; Gaps 3;

Qy 1 MNRGVPFRHLILVQLALIPATQGNKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
Db 1 MNRGVPFRHLILVQLALIPATQGNKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
Qy 61 ILGNQGSFLLTGKSPKLNDRADSRSLWDQGNFPLIINKLKIESDPTVICEVEDQKEEVOL 120
Db 61 ILGNQGSFLLTGKSPKLNDRADSRSLWDQGNFPLIINKLKIESDPTVICEVEDQKEEVOL 120
Qy 121 LVFGLTANSDTHLQOGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLSSVQLELQDSG 180
Db 121 LVFGLTANSDTHLQOGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLSSVQLELQDSG 180
Qy 181 TWCTCTVLQNGKQKVEFKIDIVLAFQKASSIYKKEGEVSFPPLA 226
Db 181 TWCTCTVLQNGKQKVEFKIDIVLAF--AST-----KGPSV---FPLA 216

RESULT 48
US-09-409-006A-4
Sequence 4, Application US/09409006A
Patent No. 6342586
GENERAL INFORMATION:
APPLICANT: Progenics Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/409,006A
FILING DATE: 29-SEP-1999
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-09-409-006A-4

Query Match 45.0%; Score 1046; DB 4; Length 530;
Best Local Similarity 93.4%; Pred. No. 3,6e-79;
Matches 211; Conservative 2; Mismatches 3; Indels 10; Gaps 3;

Qy 1 MNRGVPFRHLILVQLALIPATQGNKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
Db 1 MNRGVPFRHLILVQLALIPATQGNKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
Qy 61 ILGNQGSFLLTGKSPKLNDRADSRSLWDQGNFPLIINKLKIESDPTVICEVEDQKEEVOL 120
Db 61 ILGNQGSFLLTGKSPKLNDRADSRSLWDQGNFPLIINKLKIESDPTVICEVEDQKEEVOL 120
Qy 121 LVFGLTANSDTHLQOGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLSSVQLELQDSG 180
Db 121 LVFGLTANSDTHLQOGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLSSVQLELQDSG 180
Qy 181 TWCTCTVLQNGKQKVEFKIDIVLAFQKASSIYKKEGEVSFPPLA 226
Db 181 TWCTCTVLQNGKQKVEFKIDIVLAF--AST-----KGPSV---FPLA 216

RESULT 49
US-08-484-681-4
Sequence 4, Application US/08484681

```
Patent No. 6451313
GENERAL INFORMATION:
APPLICANT: Beaudry, Gary A.
TITLE OF INVENTION: CD4-GAMMA2 CD4-IgG2 CHIMERAS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,681
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 37690-II-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-08-484-681-4

Query Match          45.0%; Score 1046; DB 4; Length 530;
Best Local Similarity 93.4%; Pred. No. 3,66-79;
Matches 21; Conservative 2; Mismatches 3; Indels 10; Gaps 3;

QY 1 MNRGVPRFRLHLVQLALPPATQGNKVLGKKGDTVELTCTASQKSIQPHMKNNOIK 60
DB 1 MNRGVPRFRLHLVQLALPPATQGNKVLGKKGDTVELTCTASQKSIQPHMKNNOIK 60
QY 61 ILNGGSFLTKGPKSLNDRADSRSLMDQGNFPLIIKIKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILNGGSFLTKGPKSLNDRADSRSLMDQGNFPLIIKIKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDPHTLLOQGSLLTLESPPSSPVOCRSRPGKNIOGKTLVSQLELDOSG 180
DB 121 LVFGLTANSDPHTLLOQGSLLTLESPPSSPVOCRSRPGKNIOGKTLVSQLELDOSG 180
QY 181 TWCTCTVLONQKKEFKIDIVLAFQKASIVYKKEGEQVEFSFPLA 226
DB 181 TWCTCTVLONQKKEFKIDIVLAF--AST-----KGPSV---FPLA 216

RESULT 50
PCT-US93-07422-4
Sequence 4, Application PC/TUS9307422
GENERAL INFORMATION:
APPLICANT: Progenics Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: Cooper & Dunham
```

```
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07422
FILING DATE: 19930806
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
PCT-US93-07422-4

Query Match          45.0%; Score 1046; DB 5; Length 530;
Best Local Similarity 93.4%; Pred. No. 3,66-79;
Matches 21; Conservative 2; Mismatches 3; Indels 10; Gaps 3;

QY 1 MNRGVPRFRLHLVQLALPPATQGNKVLGKKGDTVELTCTASQKSIQPHMKNNOIK 60
DB 1 MNRGVPRFRLHLVQLALPPATQGNKVLGKKGDTVELTCTASQKSIQPHMKNNOIK 60
QY 61 ILNGGSFLTKGPKSLNDRADSRSLMDQGNFPLIIKIKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILNGGSFLTKGPKSLNDRADSRSLMDQGNFPLIIKIKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDPHTLLOQGSLLTLESPPSSPVOCRSRPGKNIOGKTLVSQLELDOSG 180
DB 121 LVFGLTANSDPHTLLOQGSLLTLESPPSSPVOCRSRPGKNIOGKTLVSQLELDOSG 180
QY 181 TWCTCTVLONQKKEFKIDIVLAFQKASIVYKKEGEQVEFSFPLA 226
DB 181 TWCTCTVLONQKKEFKIDIVLAF--AST-----KGPSV---FPLA 216

RESULT 51
US-08-477-4608-2
Sequence 2, Application US/084774608
Patent No. 6034223
GENERAL INFORMATION:
APPLICANT: Progenics Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
```

ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,460B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/ATM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-08-477-460B-2

Query Match 44.9%; Score 1045; DB 3; Length 432;
Best Local Similarity 98.5%; Pred. No. 3.2e-79;
Matches 203; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPPATOGKVVLGKGDVVELCTASQKKSIOFHWKNSQIK 60
DB 1 MNRGVPFRHLVLVQLALPPATOGKVVLGKGDVVELCTASQKKSIOFHWKNSQIK 60
QY 61 ILNGQSFLLTGKPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILNGQSFLLTGKPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTVLQONQKVEFKIDIVLAFQK 206
DB 181 TWCTVLQONQKVEFKIDIVLAFER 206

RESULT 52
US-08-379-516-2
Sequence 2, Application US/08379516
Patent No. 6083478
GENERAL INFORMATION:
APPLICANT: Allaway, Graham P.
APPLICANT: Madden, Paul J.
TITLE OF INVENTION: No. 6083478-Peptide1 Moieity-Conjugated CD4-Gamma2 and CD4-IgG2
TITLE OF INVENTION: Immunocjugates and Uses Thereof
FILE REFERENCE: 41215-A-PCT-US
CURRENT APPLICATION NUMBER: US/08/379,516
CURRENT FILING DATE: 1996-06-10
EARLIER APPLICATION NUMBER: PCT/US93/07422
EARLIER FILING DATE: 1993-08-06
EARLIER APPLICATION NUMBER: 07/927,931
EARLIER FILING DATE: 1992-08-07
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2

LENGTH: 432
TYPE: PRT
ORGANISM: Homo sapiens
US-08-379-516-2

Query Match 44.9%; Score 1045; DB 3; Length 432;
Best Local Similarity 98.5%; Pred. No. 3.2e-79;
Matches 203; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPPATOGKVVLGKGDVVELCTASQKKSIOFHWKNSQIK 60
DB 1 MNRGVPFRHLVLVQLALPPATOGKVVLGKGDVVELCTASQKKSIOFHWKNSQIK 60
QY 61 ILNGQSFLLTGKPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILNGQSFLLTGKPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTVLQONQKVEFKIDIVLAFQK 206
DB 181 TWCTVLQONQKVEFKIDIVLAFER 206

RESULT 53
US-09-329-916-2
Sequence 2, Application US/09329916
Patent No. 6177549
GENERAL INFORMATION:
APPLICANT: Progenics Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/329,916
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/477,460
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/ATM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: homo sapien

CELL TYPE: lymphocyte
US-09-329-916-2

Query Match 44.9%; Score 1045; DB 3; Length 432;

Best Local Similarity 98.5%; Pred. No. 3.2e-79; Mismatches 2; Indels 0; Gaps 0;

Matches 203; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPAATQGNKVVLGKKGDTVELTCTASQKSIQPHMKNNOIK 60
DB 1 MNRGVPFRHLVLVQLALPAATQGNKVVLGKKGDTVELTCTASQKSIQPHMKNNOIK 60
QY 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNPPLIKLKIETSDTYICEVEDQKEEYQL 120
DB 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNPPLIKLKIETSDTYICEVEDQKEEYQL 120
QY 121 LVFGILTANSDTHLQGGSLTTLSPSGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGILTANSDTHLQGGSLTTLSPSGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLQNKVKVEFKIDIVLAFOK 206
DB 181 TWTCTVLQNKVKVEFKIDIVLAFFR 206

RESULT 54

US-08-485-372A-2
Sequence 2, Application US/08485372A

Patent No. 6187748
GENERAL INFORMATION:

APPLICANT: Beauty, Gary A.

APPLICANT: Maddon, Paul J.

TITLE OF INVENTION: CD4-GAMMA2 CD4-IGG2 CHIMERAS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485.372A

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,227

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 37690-II-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 432 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

ORGANISM: homo sapien

CELL TYPE: lymphocyte

US-08-485-372A-2

Query Match 44.9%; Score 1045; DB 3; Length 432;

Best Local Similarity 98.5%; Pred. No. 3.2e-79; Mismatches 2; Indels 0; Gaps 0;

Matches 203; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPAATQGNKVVLGKKGDTVELTCTASQKSIQPHMKNNOIK 60
DB 1 MNRGVPFRHLVLVQLALPAATQGNKVVLGKKGDTVELTCTASQKSIQPHMKNNOIK 60
QY 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNPPLIKLKIETSDTYICEVEDQKEEYQL 120
DB 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNPPLIKLKIETSDTYICEVEDQKEEYQL 120
QY 121 LVFGILTANSDTHLQGGSLTTLSPSGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGILTANSDTHLQGGSLTTLSPSGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLQNKVKVEFKIDIVLAFOK 206
DB 181 TWTCTVLQNKVKVEFKIDIVLAFFR 206

RESULT 55

US-09-409-006A-2
Sequence 2, Application US/09409006A

Patent No. 6342586
GENERAL INFORMATION:

APPLICANT: Progenics Pharmaceuticals, Inc.

TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED

TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IGG2 IMMUNOCONJUGATES, AND USES THEREOF

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Cooper & Dunham

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/409,006A

FILING DATE: 29-SEP-1999

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/927,931

FILING DATE: 07-AUG-1992

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/AJM

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550

TELEFAX: (212) 977-9809

TELEX: 422523 COOP UT

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 432 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

ORGANISM: homo sapien

CELL TYPE: lymphocyte

US-09-409-006A-2

Query Match 44.9%; Score 1045; DB 4; Length 432;
Best Local Similarity 98.5%; Pred. No. 3.2e-79; Mismatches 2; Indels 0; Gaps 0;
Matches 203; Conservative 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPAATQGNKVVLGKKGDTVELTCTASQKSIQPHMKNNOIK 60

Db 1 MNRGVPFRHLLVLTQALLPATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNSNQIK 60
Qy 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIETSDTYICEVEDQKEVQL 120
Db 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIETSDTYICEVEDQKEVQL 120
Qy 121 LVFGLTANSDPHTLLQGOSLTLTLSPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
Db 121 LVFGLTANSDPHTLLQGOSLTLTLSPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
Qy 181 TWCTVVLQONOKVEFKIDIVLAFQK 206
Db 181 TWCTVVLQONOKVEFKIDIVLAFER 206

RESULT 56
US-08-484-681-2
; Sequence 2, Application US/08484681
; Patent No. 6451313
; GENERAL INFORMATION:
; APPLICANT: Beaudry, Gary A.
; APPLICANT: Maddon, Paul J.
; TITLE OF INVENTION: CD4-GAMMA2 CD4-IgG2 CHIMERAS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,681
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 37690-II-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: homo sapien
; CELL TYPE: lymphocyte
; US-08-484-681-2

Query Match 44.9%; Score 1045; DB 4; Length 432;
Best Local Similarity 98.5%; Pred. No. 3.2e-79;
Matches 203; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MNRGVPFRHLLVLTQALLPATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNSNQIK 60
Db 1 MNRGVPFRHLLVLTQALLPATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNSNQIK 60
Qy 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIETSDTYICEVEDQKEVQL 120
Db 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIETSDTYICEVEDQKEVQL 120
Qy 121 LVFGLTANSDPHTLLQGOSLTLTLSPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
Db 121 LVFGLTANSDPHTLLQGOSLTLTLSPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
Qy 181 TWCTVVLQONOKVEFKIDIVLAFQK 206
Db 181 TWCTVVLQONOKVEFKIDIVLAFER 206

Qy 121 LVFGLTANSDPHTLLQGOSLTLTLSPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
Db 121 LVFGLTANSDPHTLLQGOSLTLTLSPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
Qy 181 TWCTVVLQONOKVEFKIDIVLAFQK 206
Db 181 TWCTVVLQONOKVEFKIDIVLAFER 206

RESULT 57
PCT-US93-07422-2
; Sequence 2, Application PC/TUS9307422
; GENERAL INFORMATION:
; APPLICANT: Progenics Pharmaceuticals, Inc.
; TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
; TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IgG2 IMMUNOCONJUGATES, AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07422
; FILING DATE: 19930806
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/927,931
; FILING DATE: 07-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 977-9809
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: homo sapien
; CELL TYPE: lymphocyte
; PCT-US93-07422-2

Query Match 44.9%; Score 1045; DB 5; Length 432;
Best Local Similarity 98.5%; Pred. No. 3.2e-79;
Matches 203; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MNRGVPFRHLLVLTQALLPATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNSNQIK 60
Db 1 MNRGVPFRHLLVLTQALLPATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNSNQIK 60
Qy 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIETSDTYICEVEDQKEVQL 120
Db 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIETSDTYICEVEDQKEVQL 120
Qy 121 LVFGLTANSDPHTLLQGOSLTLTLSPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
Db 121 LVFGLTANSDPHTLLQGOSLTLTLSPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
Qy 181 TWCTVVLQONOKVEFKIDIVLAFQK 206
Db 181 TWCTVVLQONOKVEFKIDIVLAFER 206

Db 181 TWTCTVLGNQKKVEFKIDIVLA 206

```
|||||
RESULT 58
US-08-284-391B-31
/ Sequence 31, Application US/08284391B
/ Patent No. 5851828
/ GENERAL INFORMATION:
/ APPLICANT: Seed, Brian
/ APPLICANT: Banapour, Babak
/ APPLICANT: Romeo, Charles
/ APPLICANT: Kolanus, Waldemar
/ TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
/ TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS
/ NUMBER OF SEQUENCES: 53
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Clark & Elbing LLP
/ STREET: 176 Federal Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02110
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/284,391B
/ FILING DATE: 02-AUG-1994
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/195,395
/ FILING DATE: 14-FEB-1994
/ APPLICATION NUMBER: 07/847,566
/ FILING DATE: 06-MAR-1992
/ APPLICATION NUMBER: 07/665,961
/ FILING DATE: 07-MAR-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Elbing, Karen L.
/ REGISTRATION NUMBER: 35,238
/ REFERENCE/DOCKET NUMBER: 00786/247001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-428-0200
/ TELEFAX: 617-428-7045
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 31:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 203 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-284-391B-31

Query Match 44.8%; Score 1041; DB 2; Length 203;
Best Local Similarity 100.0%; Pred. No. 2.3e-79;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPAPATQGNKVVLGKKGDTVELTCTASQKKSIOFHMKNSNOIK 60
DB 1 MNRGVPFRHLVLVQLALPAPATQGNKVVLGKKGDTVELTCTASQKKSIOFHMKNSNOIK 60
QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKLIKIEDSDTYICEVEDQKEVQL 120
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKLIKIEDSDTYICEVEDQKEVQL 120
QY 121 LVFGILTANSDTHLLQGGSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGILTANSDTHLLQGGSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 121 LVFGILTANSDTHLLQGGSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGILTANSDTHLLQGGSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLGNQKKVEFKIDIVLA 203
```

Db 181 TWTCTVLGNQKKVEFKIDIVLA 203

```
|||||
RESULT 59
US-09-218-950-31
/ Sequence 31, Application US/09218950
/ Patent No. 6284240
/ GENERAL INFORMATION:
/ APPLICANT: Seed, Brian
/ APPLICANT: Banapour, Babak
/ APPLICANT: Romeo, Charles
/ APPLICANT: Kolanus, Waldemar
/ TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
/ TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS
/ NUMBER OF SEQUENCES: 53
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Clark & Elbing LLP
/ STREET: 176 Federal Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02110
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/218,950
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/284,391
/ FILING DATE: 02-AUG-1994
/ APPLICATION NUMBER: 08/195,395
/ FILING DATE: 14-FEB-1994
/ APPLICATION NUMBER: 07/847,566
/ FILING DATE: 06-MAR-1992
/ APPLICATION NUMBER: 07/665,961
/ FILING DATE: 07-MAR-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Elbing, Karen L.
/ REGISTRATION NUMBER: 35,238
/ REFERENCE/DOCKET NUMBER: 00786/247001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-428-0200
/ TELEFAX: 617-428-7045
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 31:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 203 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-09-218-950-31

Query Match 44.8%; Score 1041; DB 3; Length 203;
Best Local Similarity 100.0%; Pred. No. 2.3e-79;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPAPATQGNKVVLGKKGDTVELTCTASQKKSIOFHMKNSNOIK 60
DB 1 MNRGVPFRHLVLVQLALPAPATQGNKVVLGKKGDTVELTCTASQKKSIOFHMKNSNOIK 60
QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKLIKIEDSDTYICEVEDQKEVQL 120
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKLIKIEDSDTYICEVEDQKEVQL 120
QY 121 LVFGILTANSDTHLLQGGSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGILTANSDTHLLQGGSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 121 LVFGILTANSDTHLLQGGSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGILTANSDTHLLQGGSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
```


| | | | |
|----|-----|--|-----|
| Qy | 113 | QLVFRGLTANSDFHLLQGGSLTLTLLES-PPGSSPSVQCHSPRGKATIQGGKTLSSVLSQLEIQ | 177 |
| Dh | 121 | ELWVFRFTENPGRLLQGGSLTLTLIDSNPVSOPPIECCKKSSNIYKDSNAFSTHSLRIQ | 180 |
| Qy | 178 | DSGTMTCSTVLQONQKVEFEKIDIVLWFAQKASSIVYKKEGQVEFSPPLAFVTEKLTGSGE | 237 |
| Dh | 181 | DSGIMNCTVTLLNQKKHSFDMKLSTLQAGASISITAYKSEBSNAFSPPLNLGESSL--QGE | 238 |
| Qy | 238 | LWMOAEASASSKSKSITITFDLNQKESVSKRVITQDPKLNQKKLLPLHLTLPOALPOYAGSGUL | 297 |
| Dh | 239 | LWMAEKAEPSSQSITITSLNKOKRVSQKSTSNRPQLSEPLTLPIQVSIQAPAGSGUL | 298 |
| Qy | 298 | TLALEATQGLKHQEVNLVVMRATQLOLN-LTCEWVGPTSPKMLSLKLENKAKYSIKREK | 356 |
| Dh | 299 | TLTLD--KGLIYQEVNLVVMKVTQDPBSNTLTCEWMTSPSKMLIKENQEARYSRQEK | 356 |
| Qy | 357 | PWVVLNPEAGMOCCLSDSGQVLLIESNIKVL | 387 |
| Dh | 357 | VIOVAEPAGVMOCCLSEGEVVMDSKIQVL | 387 |

RESULT 62
US-08-394-442B-8
; Sequence 8, Application US/08394442B

```

1 GENERAL INFORMATION:
2 APPLICANT: Herkend, Thierry
3 APPLICANT: Tietbel, Frederic
4 TITLE OF INVENTION: New Proteins Produced By Human
5 TITLE OF INVENTION: Lymphocytes, DNA Sequences Encoding These Proteins And
6 TITLE OF INVENTION: Their Pharmaceutical And Biological Uses
7 NUMBER OF SEQUENCES: 11
8 CORRESPONDENCE ADDRESS:
9 ADDRESSEE: BROWDY AND NEWMARK, P.L.L.C.
10 STREET: 419 Seventh Street, N.W., Suite 400
11 CITY: Washington
12 STATE: D.C.
13 COUNTRY: U.S.A.
14 ZIP: 20004
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: Patent In Release #1.0, Version #1.25
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/06/394,442B
22 FILING DATE: 24-FEB-1995
23 CLASSIFICATION: 424
24 ATTORNEY/AGENT INFORMATION:
25 NAME: BROWDY, Roger L.
26 REGISTRATION NUMBER: 25,618
27 REFERENCE/DOCKET NUMBER: HERKEND=2A
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: (202) 628-5197
30 TELEFAX: (202) 737-3528
31 INFORMATION FOR SEQ ID NO: 8:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 457 amino acids
34 TYPE: amino acid
35 STRANDEDNESS: single
36 TOPOLOGY: linear
37 MOLECULE TYPE: protein
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
```

| | | | | |
|-----------------------|------------------|--------------------|-----------|-------------|
| Query Match | 43.0% | Score 999; | DB 2; | Length 457; |
| Best Local Similarity | 52.9%; | Pred. No. 2.5e-75; | | |
| Matches 207; | Conservative 61; | Mismatches 115; | Indels 8; | Gaps 5; |

```
Oy      1 MNRGVPFRHL--LLVLQLALLPAATQGNKVKLGKKDYTELTCTASQKSIQFHWNISNQ   58
        ||| |||| :|||: ||| |||||: ||| :||| :|||: |||: |
db      1 MCRGFSEFRRHLLP LLLQLSKLVVTOGTAVLGEKGSALPCESTSRRSASFAMWSSDO   60
```

[illegible]

RESULT 63
US-08-630-172-17
: Sequence 17, Application US/08630172
: Patent No. 6060054

APPLICANT: Staerz, Uwe
 TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
 TITLE OF INVENTION: LYMPHOCYTE VETO
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sheridan Ross & McIntosh
 STREET: 1700 Lincoln Street, 35th Floor
 CITY: Denver
 STATE: Colorado
 COUNTRY: U.S.
 ZIP: 80203
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/650,172
 FILING DATE:
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.
 REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2879-36
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ. ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 410 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-630-172-17

| | | | | |
|---------------------------|-------|--------------------|------------|-------------|
| Query Match | 39.2% | Score 912.5; | DB 3; | Length 410; |
| Best Local Similarity | 55.2% | Pred. No. 3.7e-68; | | |
| Matches 208; Conservative | 27; | Mismatches 71; | Indels 71; | Gaps 8; |

QY 26 NKVLGKGDVELTCTASÖKKSIOFHWNNOIKILNÖGGSFLTGPSPKLNDRADSRRS 85
 Db 1 NKVLGKGDVELTCTASÖKKSIOFHWNNOIKILNÖGGSFLTGPSPKLNDRADSRRS 60

QY 86 LMDGNPFLIIKNIKIEDSTYICEVEDOKEEVOLLVFGLTANSDFHLLQGOSLTLTLES 145
DB 61 LMDGNPFLIIKNIKIEDSTYICEVEDOKEEVOLLVFGLTANSDFHLLQGOSLTLTLES 120
QY 146 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTCTVLQONOKKVEFKIDIVLAFQ 205
DB 121 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTCTVLQONOKKVEFKIDIVLAFQ 178
QY 206 KASSIYKKEGEVEFSPLAFTVEKLTGSGELMWQAEARASSKSWITFDLKNKEVSVKR 265
DB 179 -----EPRGPTIKPCPCPCAPNLLG-----GPSVFIFPPKIKDVLMI 218
QY 266 VT-----QDPKQMG---KKLPHLTLPLALPOYAGSGNLTALAEKTKLHQ 310
DB 219 LSPITVCVVVDSEDDPDVQISWFFVNNVEVH-----TAQTQTHREHY 260
QY 311 EVNLVVRATQLO-----KNLTCEVWGPTSPKMLSLKENKEAKVSRKRPVWVLP 363
DB 261 NSRLRVVSALPIQHODMWSGKEFKCKV---NNKDLPAPIERTISKPGSVRAPQVYVLP 317
QY 364 EAGMWQ-----CLSD 374
DB 318 PEEMTKKQVTLTCMVTD 334

RESULT 64
US-09-375-419-17
; Sequence 17, Application US/09375419
; Patent No. 6264850
; GENERAL INFORMATION:
; APPLICANT: Staerz, Uwe
; TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
; TITLE OF INVENTION: LYMPHOCYTE VETO
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, 35th Floor
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/375,419
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/630,172
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2879-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-375-419-17

Query Match 39.2%; Score 912.5; DB 3; Length 410;
Best Local Similarity 55.2%; Pred. No. 3,76-68;
Matches 208; Conservative 27; Mismatches 71; Indels 71; Gaps 8;

QY 26 NKVVLGKKGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTQGPSKLNDRADSRRS 85

DB 1 NKVVLGKKGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTQGPSKLNDRADSRRS 60
QY 86 LMDGNPFLIIKNIKIEDSTYICEVEDOKEEVOLLVFGLTANSDFHLLQGOSLTLTLES 145
DB 61 LMDGNPFLIIKNIKIEDSTYICEVEDOKEEVOLLVFGLTANSDFHLLQGOSLTLTLES 120
QY 146 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTCTVLQONOKKVEFKIDIVLAFQ 205
DB 121 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTCTVLQONOKKVEFKIDIVLAFQ 178
QY 206 KASSIYKKEGEVEFSPLAFTVEKLTGSGELMWQAEARASSKSWITFDLKNKEVSVKR 265
DB 179 -----EPRGPTIKPCPCPCAPNLLG-----GPSVFIFPPKIKDVLMI 218
QY 266 VT-----QDPKQMG---KKLPHLTLPLALPOYAGSGNLTALAEKTKLHQ 310
DB 219 LSPITVCVVVDSEDDPDVQISWFFVNNVEVH-----TAQTQTHREHY 260
QY 311 EVNLVVRATQLO-----KNLTCEVWGPTSPKMLSLKENKEAKVSRKRPVWVLP 363
DB 261 NSRLRVVSALPIQHODMWSGKEFKCKV---NNKDLPAPIERTISKPGSVRAPQVYVLP 317
QY 364 EAGMWQ-----CLSD 374
DB 318 PEEMTKKQVTLTCMVTD 334

RESULT 65
US-08-021-601-12
; Sequence 12, Application US/08021601
; Patent No. 5591631
; GENERAL INFORMATION:
; APPLICANT: Leppa, Stephen H.
; APPLICANT: Kimpel, Kurt R.
; APPLICANT: Nichols, Peter J.
; APPLICANT: Atora, Naveen
; APPLICANT: Singh, Yogendra
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Needle & Rosenberg, P.C.
; STREET: 133 Carnegie Way, Suite 400
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,601
; FILING DATE: 19930212
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwendolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414,057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-021-601-12

Query Match 39.2%; Score 911; DB 1; Length 903;

Best Local Similarity 100.0%; Pred. No. 1.6e-67;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KVLGKGGTVELTCTASQKSIQFHWKNSNQIKILGNSGFLTKGPKLNDRADSRSL 86
DB 727 KVLGKGGTVELTCTASQKSIQFHWKNSNQIKILGNSGFLTKGPKLNDRADSRSL 786

QY 87 WDOGNFPLIKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDPHLLQGOSLTLTLESP 146
DB 787 WDOGNFPLIKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDPHLLQGOSLTLTLESP 846

QY 147 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTWTCTVLQNKVFEFKIDIVLA 203
DB 847 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTWTCTVLQNKVFEFKIDIVLA 903

RESULT 66
US-08-082-849B-12
; Sequence 12, Application US/08082849B
; Patent No. 5677274
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Kimpel, Kurt R.
; APPLICANT: Aroza, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
; TITLE OF INVENTION: Related Methods
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,849B
; FILING DATE: 25-JUN-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,601
; FILING DATE: 12-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-161-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-082-849B-12

Query Match 39.2%; Score 911; DB 1; Length 903;
Best Local Similarity 100.0%; Pred. No. 1.6e-67;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KVLGKGGTVELTCTASQKSIQFHWKNSNQIKILGNSGFLTKGPKLNDRADSRSL 86
DB 727 KVLGKGGTVELTCTASQKSIQFHWKNSNQIKILGNSGFLTKGPKLNDRADSRSL 786

QY 87 WDOGNFPLIKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDPHLLQGOSLTLTLESP 146
DB 847 WDOGNFPLIKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDPHLLQGOSLTLTLESP 146

DB 787 WDOGNFPLIKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDPHLLQGOSLTLTLESP 846
QY 147 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTWTCTVLQNKVFEFKIDIVLA 203
DB 847 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTWTCTVLQNKVFEFKIDIVLA 903

RESULT 67
PCT-US94-01624-12
; Sequence 12, Application PC/TUS9401624
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Kimpel, Kurt R.
; APPLICANT: Aroza, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: TOWNSEND AND TOWNSEND KHOURIE AND CREW
; STREET: Steuart Street Tower, 20th Floor, One Market
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01624
; FILING DATE: June 25, 1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-01624-12

Query Match 39.2%; Score 911; DB 5; Length 903;
Best Local Similarity 100.0%; Pred. No. 1.6e-67;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KVLGKGGTVELTCTASQKSIQFHWKNSNQIKILGNSGFLTKGPKLNDRADSRSL 86
DB 727 KVLGKGGTVELTCTASQKSIQFHWKNSNQIKILGNSGFLTKGPKLNDRADSRSL 786

QY 87 WDOGNFPLIKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDPHLLQGOSLTLTLESP 146
DB 787 WDOGNFPLIKLKIEDSDTYICEVEDQKEEVQLVFGLTANSDPHLLQGOSLTLTLESP 846

QY 147 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTWTCTVLQNKVFEFKIDIVLA 203
DB 847 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTWTCTVLQNKVFEFKIDIVLA 903

RESULT 68
US-08-630-172-1
; Sequence 1, Application US/08630172
; Patent No. 6060054
; GENERAL INFORMATION:

APPLICANT: Staerz, Uwe
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
TITLE OF INVENTION: LYMPHOCYTE VETO
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, 35th Floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,172
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-36
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-630-172-1

Query Match 30.8%; Score 716; DB 3; Length 138;
Best Local Similarity 100.0%; Pred. No. 2e-52;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 NKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIKILNGSFLTKGPSKLNDRADSRRS 85
DB 1 NKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIKILNGSFLTKGPSKLNDRADSRRS 60
QY 86 LMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVLVFGLTANSPTHLLOGSILTTLES 145
DB 61 LMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVLVFGLTANSPTHLLOGSILTTLES 120
QY 146 PGSSPSVQCRSPRGKNI 163
DB 121 PGSSPSVQCRSPRGKNI 138

RESULT 69
US-09-375-419-1
Sequence 1, Application US/09375419
Patent No. 6264950
GENERAL INFORMATION:
APPLICANT: Staerz, Uwe
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
TITLE OF INVENTION: LYMPHOCYTE VETO
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, 35th Floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/375,419
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,172
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-36
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-375-419-1

Query Match 30.8%; Score 716; DB 3; Length 138;
Best Local Similarity 100.0%; Pred. No. 2e-52;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 NKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIKILNGSFLTKGPSKLNDRADSRRS 85
DB 1 NKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIKILNGSFLTKGPSKLNDRADSRRS 60
QY 86 LMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVLVFGLTANSPTHLLOGSILTTLES 145
DB 61 LMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVLVFGLTANSPTHLLOGSILTTLES 120
QY 146 PGSSPSVQCRSPRGKNI 163
DB 121 PGSSPSVQCRSPRGKNI 138

RESULT 70
US-08-558-269-10
Sequence 10, Application US/08558269
Patent No. 5961973
GENERAL INFORMATION:
APPLICANT: Crea, Roberto
TITLE OF INVENTION: PATHOGEN-TARGETED BIOCATALYSTS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/558,269
FILING DATE: 13-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,800
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CRT-001CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-558-269-10

Query Match 25.9%; Score 603; DB 2; Length 376;
Best Local Similarity 46.6%; Pred. No. 2.5e-42;
Matches 159; Conservative 26; Mismatches 96; Indels 60; Gaps 11;

QY 27 KVLKGGKDTVELTCTASQKSIQPHMKNSNOIKILGQSGFLITGSPSKLNDRADSRSL 86
DB 3 KVLGKKGDTELTCTASQKSIQPHMKNSNOIKILGQSGFLITGSPSKLNDRADSRSL 62
QY 87 WDQGNFPLIINKLKIEDSDTYICEVEDQKEVQLVFGLTANSPTHLHQSGSLTLTLBSP 146
DB 63 WDQGNFPLIINKLKIEDSDTYICEVEDQKEVQLVFGLTANSPTHLHQSGSLTLTLBGS 122
QY 147 P---GSSP-SVQCRSPRGKNIQGGKTLVSQLELDGSGTWCT---VLQNKVVEPKID 198
DB 123 DAELGMSFWQVWLFRKSPQELLCGASLI-----SDRWVLTAAHCILYPPMDKNFTEN 174
QY 199 IYVLAFOKASSIVYKKEGEVFFPLAFTVEKLTGSGELMWQARASSKSWITFDLKN 258
DB 175 DLVVGKHSRTYRERNIETISM-----LEKTIYHPRYVW---RENLDRIALMKLK- 223
QY 259 KEVSVK-----RVTDPKLQMGKKLPLHLTLPOLPOYAGSGNLTLLAEKATGK- 307
DB 224 KPVAFSDYIHVPCLPDRETAASLLQAGYK-----GRVTGNGWLKXTWTWNVGKG 272
QY 308 ---LHGEVNLVYMR-----ATQLQKNLTCEVWGPTSPK 337
DB 273 QPSVLQVYNLPIYERPVCKDSTRIRITDMFCAGYKPDGK 313

RESULT 71

US-09-410-882-10
Sequence 10, Application US/09410882
Patent No. 6287561
GENERAL INFORMATION:
APPLICANT: Crea, Roberto
TITLE OF INVENTION: PATHOGEN-TARGETED BIOCATALYSTS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/410,882
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/558,269
FILING DATE: 13-NOV-1995
APPLICATION NUMBER: US 07/847,800
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CRI-001CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-410-882-10

Query Match 25.9%; Score 603; DB 3; Length 376;
Best Local Similarity 46.6%; Pred. No. 2.5e-42;
Matches 159; Conservative 26; Mismatches 96; Indels 60; Gaps 11;

QY 27 KVLKGGKDTVELTCTASQKSIQPHMKNSNOIKILGQSGFLITGSPSKLNDRADSRSL 86
DB 3 KVLGKKGDTELTCTASQKSIQPHMKNSNOIKILGQSGFLITGSPSKLNDRADSRSL 62
QY 87 WDQGNFPLIINKLKIEDSDTYICEVEDQKEVQLVFGLTANSPTHLHQSGSLTLTLBSP 146
DB 63 WDQGNFPLIINKLKIEDSDTYICEVEDQKEVQLVFGLTANSPTHLHQSGSLTLTLBGS 122
QY 147 P---GSSP-SVQCRSPRGKNIQGGKTLVSQLELDGSGTWCT---VLQNKVVEPKID 198
DB 123 DAELGMSFWQVWLFRKSPQELLCGASLI-----SDRWVLTAAHCILYPPMDKNFTEN 174
QY 199 IYVLAFOKASSIVYKKEGEVFFPLAFTVEKLTGSGELMWQARASSKSWITFDLKN 258
DB 175 DLVVGKHSRTYRERNIETISM-----LEKTIYHPRYVW---RENLDRIALMKLK- 223
QY 259 KEVSVK-----RVTDPKLQMGKKLPLHLTLPOLPOYAGSGNLTLLAEKATGK- 307
DB 224 KPVAFSDYIHVPCLPDRETAASLLQAGYK-----GRVTGNGWLKXTWTWNVGKG 272
QY 308 ---LHGEVNLVYMR-----ATQLQKNLTCEVWGPTSPK 337
DB 273 QPSVLQVYNLPIYERPVCKDSTRIRITDMFCAGYKPDGK 313

RESULT 72

US-08-558-269-6
Sequence 6, Application US/08558269
Patent No. 5961973
GENERAL INFORMATION:
APPLICANT: Crea, Roberto
TITLE OF INVENTION: PATHOGEN-TARGETED BIOCATALYSTS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/558,269
FILING DATE: 13-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,800
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CRI-001CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 383 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-558-269-6

Query Match 25.4%; Score 589.5; DB 2; Length 383;
Best Local Similarity 72.9%; Pred. No. 3.5e-41;
Matches 121; Conservative 9; Mismatches 23; Indels 13; Gaps 2;

QY 27 KVLGKKGDVLTCTASQKKSIOFHWKNSNQIKILGNQGSFLTKGPSKLNDRADSRSL 86
DB 3 KVLGKKGDVLTCTASQKKSIOFHWKNSNQIKILGNQGSFLTKGPSKLNDRADSRSL 62
QY 87 WDOGNFPLIIKNIKIEDSDTYICEVEDQKEVOLLVFGLTANSDTHLLOGOSLTLTLESP 146
DB 63 WDOGNFPLIIKNIKIEDSDTYICEVEDQKEVOLLVFGLTANSDTHLLOGOSLTLTLESP 117
QY 147 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLQNOCK 192
DB 118 --STCGLRQYSQPQFRIGKGLFADIA-----SHPQQAIFAGKRR 155

RESULT 73
US-09-410-882-6
Sequence 6, Application US/09410882
Patent No. 6287561
GENERAL INFORMATION:
APPLICANT: Crea, Roberto
TITLE OF INVENTION: PATHOGEN-TARGETED BIOCATALYSTS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/410,882
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/558,269
FILING DATE: 13-NOV-1995
APPLICATION NUMBER: US 07/847,800
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CRI-001CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 383 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-410-882-6

Query Match 25.4%; Score 589.5; DB 3; Length 383;
Best Local Similarity 72.9%; Pred. No. 3.5e-41;
Matches 121; Conservative 9; Mismatches 23; Indels 13; Gaps 2;
QY 27 KVLGKKGDVLTCTASQKKSIOFHWKNSNQIKILGNQGSFLTKGPSKLNDRADSRSL 86
DB 3 KVLGKKGDVLTCTASQKKSIOFHWKNSNQIKILGNQGSFLTKGPSKLNDRADSRSL 62

DB 3 KVLGKKGDVLTCTASQKKSIOFHWKNSNQIKILGNQGSFLTKGPSKLNDRADSRSL 62
QY 87 WDOGNFPLIIKNIKIEDSDTYICEVEDQKEVOLLVFGLTANSDTHLLOGOSLTLTLESP 146
DB 63 WDOGNFPLIIKNIKIEDSDTYICEVEDQKEVOLLVFGLTANSDTHLLOGOSLTLTLESP 117
QY 147 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLQNOCK 192
DB 118 --STCGLRQYSQPQFRIGKGLFADIA-----SHPQQAIFAGKRR 155

RESULT 74
US-08-466-368-5
Sequence 5, Application US/08466368
Patent No. 6093539
GENERAL INFORMATION:
APPLICANT: Maddon, Paul J.
APPLICANT: Littman, Dan R.
APPLICANT: Chess, Leonard
APPLICANT: Axel, Richard
APPLICANT: Weiss, Robin
APPLICANT: McDougal, J. S.
TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN
TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,368
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 24577-EI-B/JPW/ARC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: YES
FEATURE:
NAME/KEY: Active-site
LOCATION: 1..94
US-08-466-368-5

Query Match 21.2%; Score 494; DB 3; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.2e-34;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 QGNKVVLGKKGDVLTCTASQKKSIOFHWKNSNQIKILGNQGSFLTKGPSKLNDRADSR 83
DB 1 QGNKVVLGKKGDVLTCTASQKKSIOFHWKNSNQIKILGNQGSFLTKGPSKLNDRADSR 60
QY 84 RSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKE 117
DB 61 RSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKE 94

```
RESULT 75
US-08-470-998-2
/ Sequence 2, Application US/08470998
/ Patent No. 6570000
/ GENERAL INFORMATION:
/ APPLICANT: Maddon, Paul J.
/ APPLICANT: Littman, Dan R.
/ APPLICANT: Chase, Leonard
/ APPLICANT: Axel, Richard
/ APPLICANT: Weiss, Robin
/ APPLICANT: McDougal, J. S.
/ TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN
/ TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Cooper & Dunham LLP
/ STREET: 1185 Avenue of Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/470,998
/ FILING DATE: 06-JUN-1995
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: White, John P.
/ REGISTRATION NUMBER: 28,678
/ REFERENCE/DOCKET NUMBER: 24577-F1-B/JPW/AKC
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-278-0400
/ TELEFAX: 212-391-0525
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 94 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: YES
/ ANTI-SENSE: YES
/ FEATURE:
/ NAME/KEY: Active-site
/ LOCATION: 1..94
/ US-08-470-998-2

Query Match      21.2%; Score 494; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.2e-34;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 QGNKRVLGKGDYVLTCTASQKKSIOFHWKNSNOIKILNGSGFLTGPSKLNDRASR 83
DB 1 QGNKRVLGKGDYVLTCTASQKKSIOFHWKNSNOIKILNGSGFLTGPSKLNDRASR 60

QY 84 RSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEE 117
DB 61 RSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEE 94

RESULT 76
US-08-328-500-10
/ Sequence 10, Application US/08328500
/ Patent No. 6673896
/ GENERAL INFORMATION:
/ APPLICANT: Maddon, Paul J.
/ APPLICANT: Axel, Richard
/ APPLICANT: Sweet, Richard W.
```

```
APPLICANT: Arthos, James
/ TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Cooper & Dunham
/ STREET: 1185 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/328,500
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: White, John P.
/ REGISTRATION NUMBER: 28,678
/ REFERENCE/DOCKET NUMBER: 0575/24577-CY
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 278-0400
/ TELEFAX: (212) 391-0525
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 94 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: not relevant
/ TOPOLOGY: not relevant
/ MOLECULE TYPE: protein
/ US-08-328-500-10

Query Match      21.2%; Score 494; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.2e-34;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 QGNKRVLGKGDYVLTCTASQKKSIOFHWKNSNOIKILNGSGFLTGPSKLNDRASR 83
DB 1 QGNKRVLGKGDYVLTCTASQKKSIOFHWKNSNOIKILNGSGFLTGPSKLNDRASR 60

QY 84 RSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEE 117
DB 61 RSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEE 94

RESULT 77
US-08-332-562A-84
/ Sequence 84, Application US/08332562A
/ Patent No. 5985599
/ GENERAL INFORMATION:
/ APPLICANT: MCKENZIE, Ian F.C.
/ APPLICANT: HOGARTH, Mark P.
/ APPLICANT: HIBBS, Margaret L.
/ APPLICANT: SCOTT, Bernadette M.
/ APPLICANT: BONADONNA, Lisa
/ APPLICANT: HULETT, Mark D.
/ TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
/ NUMBER OF SEQUENCES: 136
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: 3000 K Street, N.W., Suite 500
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20007-5109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,562A
; FILING DATE: 31-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/896,457
; FILING DATE: 27-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 54270/119/GRHA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-332-562A-84

```

```

Query Match 17.5%; Score 406; DB 2; Length 80;
Best Local Similarity 98.8%; Pred. No. 7.6e-27;
Matches 79; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 123 FGLTANSDTHLQOSLTLTLESPGSSPVQCRSPKXNIQGGKTLVSQLELDSDGTW 182
Db 1 FGLTANSDTHLQOSLTLTLESPGSSPVQCRSPKXNIQGGKTLVSQLELDSDGTW 60
Qy 183 TCTVLQNKQKVEFKIDIVL 202
Db 61 TCTVLQNKQKVEFKIDIVL 80

```

```

RESULT 78
US-09-050-861B-8
; Sequence 8, Application US/09050861B
; Patent No. 655314
; GENERAL INFORMATION:
; APPLICANT: Payan, Donald
; TITLE OF INVENTION: TOSO AS A TARGET FOR DRUG SCREENING
; FILE REFERENCE: RIGL-002CON
; CURRENT APPLICATION NUMBER: US/09/050,861B
; CURRENT FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: US/09/651,150B
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: US 09/050,861
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-050-861B-8

```

```

Query Match 16.9%; Score 392; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 1e-25;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 37 VELTCTASQKSIQFHMKNNSQIKILGNQGSFLTKGPKLNDRADSRRLMDQGNFPLII 96
Db 1 VELTCTASQKSIQFHMKNNSQIKILGNQGSFLTKGPKLNDRADSRRLMDQGNFPLII 60
Qy 97 KNLKIEDSDTYICE 110
Db 61 KNLKIEDSDTYICE 74

```

```

RESULT 79
US-08-332-562A-87

```

```

; Sequence 87, Application US/08332562A
; Patent No. 5985599
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, Ian F.C.
; APPLICANT: HOGARTH, Mark P.
; APPLICANT: HIBBS, Margaret L.
; APPLICANT: SCOTT, Bernadette M.
; APPLICANT: BONADONNA, Lisa
; APPLICANT: HULETT, Mark D.
; TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 136
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,562A
; FILING DATE: 31-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/896,457
; FILING DATE: 27-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 54270/119/GRHA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 87:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 83 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-332-562A-87

```

```

Query Match 16.8%; Score 391; DB 2; Length 83;
Best Local Similarity 95.2%; Pred. No. 1.4e-25;
Matches 79; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy 305 TGKLEHENVLVWRATQLOKNTLCEVWGPSPPLMTSLKLENKAKVSKREKPVWVLPPE 364
Db 1 TGKLEHENVLVWRATQLOKNTLCEVWGPSPPLMTSLKLENKAKVSKREKPVWVLPPE 60
Qy 365 AGMWQCLSDSGVLLIESNIKVL 387
Db 61 AGMWQCLSDSGVLLIESNIKVL 83

```

```

RESULT 80
US-09-100-409A-20
; Sequence 20, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York

```

STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-751-6849
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 61 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-20

Query Match 13.5%; Score 314; DB 3; Length 61;
Best Local Similarity 100.0%; Pred. No. 2.5e-19;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 SKLNRAADRSRLWDGNEPLIIKNIKIDSDPTICEVEDDKEEVLVFGILTASDTHL 133
DB 1 SKLNRAADRSRLWDGNEPLIIKNIKIDSDPTICEVEDDKEEVLVFGILTASDTHL 60
QY 134 L 134
DB 61 L 61

RESULT 81
US-08-284-391B-35
Sequence 35, Application US/08284391B
Patent No. 5651828
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566

FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/565,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-35

Query Match 12.4%; Score 288; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 3.5e-17;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 396 PRASALPAPPTGSALPDPTASALPDPPASALPALALVTSFLGLGVACVLARTR 453
DB 1 PRASALPAPPTGSALPDPTASALPDPPASALPALALVTSFLGLGVACVLARTR 58

RESULT 82
US-09-218-950-35
Sequence 35, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045

TELEX:
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-35

Query Match 12.4%; Score 288; DB 3; Length 58;
Best Local Similarity 100.0%; Pred. No. 3.5e-17;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 396 PRASALPAPPTGSALPDPTASALPDPPASALPALAVISFLGLGIVACVLART 453
Db 1 PRASALPAPPTGSALPDPTASALPDPPASALPALAVISFLGLGIVACVLART 58

RESULT 83
US-09-100-409A-10
Sequence 10, Application US/09100409A
Patent No. 6090388

GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
PREVENTION AND TREATMENT OF HIV INFECTION AND
IMMUNE DISORDERS
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-10

Query Match 9.8%; Score 228; DB 3; Length 46;
Best Local Similarity 95.6%; Pred. No. 2.6e-12;
Matches 43; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 52 HWKNSNQIKILGNQGSFLTKGPKSKLNDRADRSRLMDQGNFPLII 96
Db 2 HWKNSNQIKILGNQGSFLTKGPKSKLNDRADRSRLMDQGNFPLII 46

RESULT 84
US-09-100-409A-34
Sequence 34, Application US/09100409A

Patent No. 6090388
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
PREVENTION AND TREATMENT OF HIV INFECTION AND
IMMUNE DISORDERS
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 81 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-34

Query Match 9.8%; Score 228; DB 3; Length 81;
Best Local Similarity 95.6%; Pred. No. 6.1e-12;
Matches 43; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 52 HWKNSNQIKILGNQGSFLTKGPKSKLNDRADRSRLMDQGNFPLII 96
Db 37 HWKNSNQIKILGNQGSFLTKGPKSKLNDRADRSRLMDQGNFPLII 81

RESULT 85
US-09-100-409A-2
Sequence 2, Application US/09100409A
Patent No. 6090388
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
PREVENTION AND TREATMENT OF HIV INFECTION AND
IMMUNE DISORDERS
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A

;; FILING DATE:
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME:
;; REGISTRATION NUMBER:
;; REFERENCE/DOCKET NUMBER: 1151-4154
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-758-4800
;; TELEFAX: 212-751-6849
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 40 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-100-409A-2

Query Match 9.1%; Score 211; DB 3; Length 40;
Best Local Similarity 97.5%; Pred. No. 5.6e-11;
Matches 39; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 HKKNSNQIKILNGQSFLLTKGPSKLNDRADRSRLMDQGN 91
Db 1 HKKNNQIKILNGQSFLLTKGPSKLNDRADRSRLMDQGN 40

RESULT 86
US-09-100-409A-4
; Sequence 4, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-409A-4

Query Match 9.1%; Score 211; DB 3; Length 42;
Best Local Similarity 97.5%; Pred. No. 6e-11;
Matches 39; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 HKKNSNQIKILNGQSFLLTKGPSKLNDRADRSRLMDQGN 91

Db 2 HKKNNQIKILNGQSFLLTKGPSKLNDRADRSRLMDQGN 41

RESULT 87
US-09-100-409A-32
; Sequence 32, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-409A-32

Query Match 9.1%; Score 211; DB 3; Length 77;
Best Local Similarity 97.5%; Pred. No. 1.5e-10;
Matches 39; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 HKKNSNQIKILNGQSFLLTKGPSKLNDRADRSRLMDQGN 91
Db 37 HKKNNQIKILNGQSFLLTKGPSKLNDRADRSRLMDQGN 76

RESULT 88
5510256-5
; Patent No. 5510256
; APPLICANT: KIRSCHNER, RICHARD J., MOTT, JOHN E., ECKENRODE,
; FRANCES M., BRUNNER, DAVID P.
; TITLE OF INVENTION: ELIMINATING INTERNAL INITIATION OF
; SOLUBLE CDA GENE
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,995
; FILING DATE: 29-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 13,828
; FILING DATE: 02-FEB-1993
; APPLICATION NUMBER: 562,861
; FILING DATE: 06-AUG-1990
; SEQ ID NO: 5;
; LENGTH: 41

5510256-5

Query Match 8.6%; Score 200; DB 6; Length 41;
Best Local Similarity 97.6%; Pred. No. 4.8e-10;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 95 IIKNLIKEDSDTYICEVEDQKEEVQLVGLTANSSTHLLQ 135
Db 1 IIKNLIKEDSDTYICEVEDQKEEVQLVGLTANSSTHLLQ 41

RESULT 89

US-08-630-172-7
; Sequence 7, Application US/08630172
; Patent No. 6060054

GENERAL INFORMATION:

APPLICANT: Staerz, Uwe

TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESS: Sheridan Rose & McIntosh

STREET: 1700 Lincoln Street, 35th Floor

CITY: Denver

STATE: Colorado

COUNTRY: U.S.

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/630.172

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER: 2879-36

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 154 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-630-172-7

Query Match 7.9%; Score 184.5; DB 3; Length 154;
Best Local Similarity 57.5%; Pred. No. 7e-08;
Matches 46; Conservative 5; Mismatches 12; Indels 17; Gaps 4;

Qy 367 MWOCLSDSG---QVLEESNI-----KVLPT-----W--STVHPRASALPAPPTGSA 409
Db 75 MHRQLSDTGTTCQATEENVVYSGTLVLTBEQSGQWHRCSAPRASALPAPPTGSA 134

Qy 410 LPDPQTASALPDPASALP 429
Db 135 LPDPQTASALPDPASALP 154

RESULT 90

US-09-375-419-7

; Sequence 7, Application US/09375419

; Patent No. 6264950

GENERAL INFORMATION:

APPLICANT: Staerz, Uwe

TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T

LYMPHOCYTE VETO

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Rose & McIntosh

STREET: 1700 Lincoln Street, 35th Floor

CITY: Denver

STATE: Colorado

COUNTRY: U.S.

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/375.419

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/630,172

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER: 2879-36

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 154 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-375-419-7

Query Match 7.9%; Score 184.5; DB 3; Length 154;
Best Local Similarity 57.5%; Pred. No. 7e-08;
Matches 46; Conservative 5; Mismatches 12; Indels 17; Gaps 4;

Qy 367 MWOCLSDSG---QVLEESNI-----KVLPT-----W--STVHPRASALPAPPTGSA 409
Db 75 MHRQLSDTGTTCQATEENVVYSGTLVLTBEQSGQWHRCSAPRASALPAPPTGSA 134

Qy 410 LPDPQTASALPDPASALP 429
Db 135 LPDPQTASALPDPASALP 154

RESULT 91

US-08-332-562A-85

; Sequence 85, Application US/08332562A

; Patent No. 5985599

GENERAL INFORMATION:

APPLICANT: MCKENZIE, Ian F.C.

APPLICANT: HOGARTH, Mark P.

APPLICANT: HIBBS, Margaret L.

APPLICANT: SCOTT, Bernadette M.

APPLICANT: BONADONNA, Lisa

APPLICANT: HULETT, Mark D.

TITLE OF INVENTION: PC RECEPTOR FOR IMMUNOGLOBULIN

NUMBER OF SEQUENCES: 136

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., Suite 500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

```

1  APPLICATION NUMBER: US/08/332,562A
2  FILING DATE: 31-OCT-1994
3  CLASSIFICATION: 435
4  PRIOR APPLICATION DATA:
5  APPLICATION NUMBER: US 07/896,457
6  FILING DATE: 27-MAY-1992
7  ATTORNEY/AGENT INFORMATION:
8  NAME: BENT, Stephen A.
9  REGISTRATION NUMBER: 29,768
10 REFERENCE/DOCKET NUMBER: 54270/119/GRIA
11 TELECOMMUNICATION INFORMATION:
12 TELEPHONE: (202)672-5300
13 TELEFAX: (202)672-5359
14 TELEX: 904136
15 INFORMATION FOR SEQ ID NO: 85:
16 SEQUENCE CHARACTERISTICS:
17     LENGTH: 81 amino acids
18     TYPE: amino acid
19     STRANDEDNESS: single
20     TOPOLOGY: linear
21 US-08-332-562A-85

```

| | | | | |
|-----------------------|-------|--|-------------------------------------|-----------|
| Query Match | 7.5% | Score 174.5 | DB 2 | Length 81 |
| Best Local Similarity | 48.1% | Pred. No. 1.8e-07 | | |
| Matches | 39 | Conservative 13 | Mismatches 28 | Indels 1 |
| QY | 123 | EGLTANSTHLLQGSSTLTLES | -PPGSSPVQCRSPRGKNIQGGKITSVSGLEIDSGT | 181 |
| | | : : | : : : : : : | |
| Db | 1 | FKVFPSPETSLQGSSTLTLDNSKVNPLTECHGKKGKVVSGSKVLSMSNLRVQDSDF | | 60 |
| QY | 182 | WTCITLQKQKVEPKIDIVVL | 202 | |
| | | | : : : | |
| Db | 61 | WNCITVTLIDQKKNWFGMTSLVL | 81 | |

RESULT 92
 US-08-332-562A-86
 Sequence 86, Application US/08332562A
 Patent No. 5985599
 GENERAL INFORMATION:
 APPLICANT: MCKENZIE, Ian F.C.
 APPLICANT: HOGARTH, Mark P.
 APPLICANT: HIBBS, Margaret L.
 APPLICANT: SCOTT, Bernadette M.
 APPLICANT: BONADONNA, Lisa
 APPLICANT: HULETT, Mark D.
 TITLE OF INVENTION: PC RECEPTOR FOR IMMUNOGLOBULIN
 NUMBER OF SEQUENCES: 136
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W., Suite 500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/332.562A
 FILING DATE: 31-OCT-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/896.457
 FILING DATE: 27-MAY-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 54270/119/GRHA
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300

TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 81 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-332-562A-86

Query Match 7.5%; Score 173.5; DB 2; length 81;
Best Local Similarity 45.7%; Pred. No. 2,2e-07;
Matches 37; Conservative 10; Mismatches 33; Indels 1; Gaps 1;

DY 123 FGLTANDTHLDDGOSTLTLES -PRSSPSAVGRSRGRKNIQSGKTLVSQDLDPDST 181
DB 1 FRVFNFGSTRLLQOOSLTLLIDSNPKVDPEICKRKSNNIVKDSKAFSTHSLRIQDSGI 60

DY 182 WTCTVLONOKKVEFKIDIVYL 202
DB 61 WNCVTTLNOKKHGSFDMKLSTLYL 81

```

RESULT 93
US-09-100-409A-11
/ Sequence 11, Application US/09100409A
/ Patent No. 6090388
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Chang Yi
/ TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
/ PREVENTION AND TREATMENT OF HIV INFECTION AND
/ IMMUNE DISORDERS
/ NUMBER OF SEQUENCES: 64

```

STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentlin Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
Type: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-11

| | | | | |
|-----------------------|----------------|-------------------|----------|-----------|
| Query Match | 7.2% | Score 167 | DB 3 | Length 34 |
| Best Local Similarity | 97.0% | Pred. No. 2.1e-07 | | |
| Matches 32 | Conservative 0 | Mismatches 1 | Indels 0 | Gaps 0 |

| | | | |
|----|----|-----------------------------------|----|
| QY | 64 | NQGSFLTKGSPSKLNDRADSRRLMDQCNFPLII | 96 |
| | | | |
| Ob | 2 | NQGSFLTKGSPSKLNDRADSRRLMDQCNFPLII | 34 |
| | | | |

```
RESULT 94
US-09-100-409A-33
; Sequence 33, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-409A-33

Query Match 7.2%; Score 167; DB 3; Length 69;
Best Local Similarity 97.0%; Pred. No. 6.1e-07;
Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 64 NQGSFLTKGPKSKLNDRADSRSLMDQGNFPLII 96
Db 37 NQGSFLTKGPKSKLNDRADSRSLMDQGNCPILII 69

RESULT 95
US-09-100-409A-19
; Sequence 19, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-409A-19

Query Match 6.7%; Score 156; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 64 NQGSFLTKGPKSKLNDRADSRSLMDQGNF 92
Db 1 NQGSFLTKGPKSKLNDRADSRSLMDQGNF 29

RESULT 96
PCT-US93-00031-13
; Sequence 13, Application PC/TUS9300031
; GENERAL INFORMATION:
; APPLICANT: Osborn, Laurelee
; APPLICANT: Benjamin, Christopher D.
; TITLE OF INVENTION: ANTIBODIES RECOGNIZING THE FOURTH
; TITLE OF INVENTION: IMMUNOGLOBULIN-LIKE DOMAIN OF VCAM1
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00031
; FILING DATE: 19930112
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 92,306-A; D001 CIP PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 715-1000
; TELEFAX: (312) 715-1234
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-00031-13

Query Match 6.7%; Score 155; DB 5; Length 735;
Best Local Similarity 21.9%; Pred. No. 0.00021;
Matches 91; Conservative 71; Mismatches 166; Indels 88; Gaps 18;
```


APPLICATION NUMBER: US 07/359516
FILING DATE: 01-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/354151
FILING DATE: 28-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B124CIP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 739 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-482-073-6

Query Match 6.7%; Score 155; DB 4; Length 739;
Best Local Similarity 21.9%; Pred. No. 0.00022;
Matches 91; Conservative 71; Mismatches 166; Indels 88; Gaps 18;

8 RHLLVLTQALPLPAT--QGNKVVLGKKGDIVELTCTASQKSIOPFMKNSNQIKLGNQ 65
211 RQAVKELOVYISPKRTIVISNPSTKLOEGGSVTMTCSSEGLPAPEIFW----- 258
66 GSFPLTKGPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICE---VEDQKEVQL 121
259 -----SKLDNGNLQHL---SGNATLTLIARMEDSGIYCEGVNLIGKRKEVELI 307
122 V----FGLTANSDTHLQ--GQSLTITLESPPSSSVQCRSRGNKIQG-----GKT 168
308 VQEKPFVEISPPRIAAQIGDSVWLTCVNGCESPSFSWRTQIDPLSGVRSBGTNST 367
169 LSVSOLLEQDSGWTCTVLONOKVEFKIDIVLAFOKASSIYK---KEGEVFEFS--- 222
368 LTLSPVSENEHSYLTCTVCGHKLKEKIQVELYSFPDPRIEMSGLVNGSSVTVYSCV 427
223 --PPL-AFTVEKLTGSGELMWQERASSKSWITF---DLNKEVSVKRVTDPKLQ- 273
428 PSVYPLDRLEIELKGETIL-----ENIEFLEDITDKMSLEKNSLEMTFPIETED 476
274 MGKTL----PLHLTLPQALPOVAGSGNLTLEAKTGLHQBVLVYVMAATQLOK---N 325
477 TGAALVCOAKLHIDMEFEKQROQ---TQTLVYNVAP--RDTTVLVSPSSILEGSSVN 531
326 LTCEVWGPTSPKMLSLKLENKEAKVSKREKPVVWLNPEAGMOCCLSDSGVILE 381
532 MTLCSGFPAPKILMSRQLPNEGLOPSENATLTILSTK-----MEDSGVYLCE 580

RESULT 99
PCT-US93-00031-9
Sequence 9, Application PC/TUS9300031
GENERAL INFORMATION:

APPLICANT: Osborn, Laurelee
APPLICANT: Benjamin, Christopher D.
TITLE OF INVENTION: ANTIBODIES RECOGNIZING THE FOURTH
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Witcoff, Ltd.
STREET: 10 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: US
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00031
FILING DATE: 19930112
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 92,306-A; D001 CIP PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 715-1000
TELEFAX: (312) 715-1234
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 739 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-00031-9

Query Match 6.7%; Score 155; DB 5; Length 739;
Best Local Similarity 21.9%; Pred. No. 0.00022;
Matches 91; Conservative 71; Mismatches 166; Indels 88; Gaps 18;

8 RHLLVLTQALPLPAT--QGNKVVLGKKGDIVELTCTASQKSIOPFMKNSNQIKLGNQ 65
211 RQAVKELOVYISPKRTIVISNPSTKLOEGGSVTMTCSSEGLPAPEIFW----- 258
66 GSFPLTKGPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICE---VEDQKEVQL 121
259 -----SKLDNGNLQHL---SGNATLTLIARMEDSGIYCEGVNLIGKRKEVELI 307
122 V----FGLTANSDTHLQ--GQSLTITLESPPSSSVQCRSRGNKIQG-----GKT 168
308 VQEKPFVEISPPRIAAQIGDSVWLTCVNGCESPSFSWRTQIDPLSGVRSBGTNST 367
169 LSVSOLLEQDSGWTCTVLONOKVEFKIDIVLAFOKASSIYK---KEGEVFEFS--- 222
368 LTLSPVSENEHSYLTCTVCGHKLKEKIQVELYSFPDPRIEMSGLVNGSSVTVYSCV 427
223 --PPL-AFTVEKLTGSGELMWQERASSKSWITF---DLNKEVSVKRVTDPKLQ- 273
428 PSVYPLDRLEIELKGETIL-----ENIEFLEDITDKMSLEKNSLEMTFPIETED 476
274 MGKTL----PLHLTLPQALPOVAGSGNLTLEAKTGLHQBVLVYVMAATQLOK---N 325
477 TGAALVCOAKLHIDMEFEKQROQ---TQTLVYNVAP--RDTTVLVSPSSILEGSSVN 531
326 LTCEVWGPTSPKMLSLKLENKEAKVSKREKPVVWLNPEAGMOCCLSDSGVILE 381
532 MTLCSGFPAPKILMSRQLPNEGLOPSENATLTILSTK-----MEDSGVYLCE 580

RESULT 100
US-08-261-304-7
Sequence 7, Application US/08261304
Patent No. 5708147
GENERAL INFORMATION:

APPLICANT: Cybulsky, Myron I.
APPLICANT: Gimbrowe, Michael A.
TITLE OF INVENTION: Mononuclear Leukocyte Directed
TITLE OF INVENTION: Endothelial Adhesion Molecule Associated with
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue, N.W.
CITY: Suite 300
STATE: Washington
COUNTRY: United States of America

```

1      ZIP: 20036
2      COMPUTER READABLE FORM:
3      MEDIUM TYPE: Floppy Disk
4      COMPUTER: IBM PC compatible
5      OPERATING SYSTEM: PC-DOS/MS-DOS
6      SOFTWARE: Ascii
7      CURRENT APPLICATION DATA:
8      APPLICATION NUMBER: US/08/261,304
9      FILING DATE:
10     CLASSIFICATION: 435
11     PRIOR APPLICATION DATA:
12     APPLICATION NUMBER: 07/649,565
13     FILING DATE: 01-FEB-1991
14     APPLICATION NUMBER: U.S. 07/487,038
15     FILING DATE: 02-MAR-1990
16     ATTORNEY/AGENT INFORMATION:
17     NAME: Porter, Jane E. R.
18     REGISTRATION NUMBER: 33,332
19     REFERENCE/DOCKET NUMBER: 06227.2100004
20     TELECOMMUNICATION INFORMATION:
21     TELEPHONE: (202) 833-7533
22     INFORMATION FOR SEQ ID NO: 7:
23     SEQUENCE CHARACTERISTICS:
24     LENGTH: 662 amino acids
25     TYPE: amino acid
26     TOPOLOGY: linear
27     MOLECULE TYPE: peptide
28     /S=US-08-261-304-7

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 13:00:24 ; Search time 10.4109 Seconds
(without alignments)
4185.504 Million cell updates/sec

Title: SEOS
Perfect score: 2325
Sequence: 1 NMRGVFRRLLVLQLALP.....VISFLGLGLGVACVLRNR 453

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 125 summaries

Database :
1: PIR 78:*
2: pirl:*
3: pirl:*
4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the change being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|---------------------|
| 1 | 2032 | 87.4 | 458 | 1 | RWHT4 | T-cell surface gly |
| 2 | 1884 | 81.0 | 432 | 1 | RMCT4 | T-cell surface gly |
| 3 | 1743 | 75.0 | 432 | 1 | RMCT4 | T-cell surface gly |
| 4 | 1147 | 49.3 | 459 | 2 | A46254 | CD4 precursor - ra |
| 5 | 1101 | 47.4 | 432 | 2 | S30193 | T-cell surface gly |
| 6 | 999 | 43.0 | 457 | 2 | A27449 | T-cell surface gly |
| 7 | 993 | 42.7 | 457 | 1 | RMST4 | T-cell surface gly |
| 8 | 357 | 15.4 | 71 | 2 | I6082 | CD4 receptor - hum |
| 9 | 305.5 | 13.1 | 99 | 2 | S21461 | T-cell surface gly |
| 10 | 293 | 12.6 | 240 | 2 | A39016 | T-cell surface gly |
| 11 | 280.5 | 12.1 | 99 | 2 | S21462 | T-cell surface gly |
| 12 | 162 | 7.0 | 739 | 2 | JS0675 | vascular cell adhe |
| 13 | 155.5 | 6.7 | 538 | 2 | UC2457 | vascular cell adhe |
| 14 | 145 | 6.7 | 739 | 1 | A41288 | vascular cell adhe |
| 15 | 145 | 6.2 | 761 | 1 | IJHUNG | vascular cell adhe |
| 16 | 144 | 6.2 | 739 | 2 | JN0581 | vascular cell adhe |
| 17 | 143 | 6.2 | 647 | 2 | B41288 | vascular cell adhe |
| 18 | 138 | 5.9 | 398 | 2 | I49443 | gene 28a protein - |
| 19 | 137.5 | 5.9 | 702 | 2 | A36319 | carcinoembryonic a |
| 20 | 135 | 5.8 | 1259 | 2 | S36126 | neural cell adhesi |
| 21 | 134.5 | 5.8 | 1260 | 1 | S05479 | neural cell adhesi |
| 22 | 133 | 5.7 | 2783 | 2 | T34416 | hypothetical prote |
| 23 | 132 | 5.7 | 1011 | 2 | T13659 | neurotensin - fr |
| 24 | 132 | 5.7 | 6831 | 2 | A88852 | protein unc-22 [im |
| 25 | 132 | 5.7 | 6839 | 2 | S57242 | twitlin [similar] |
| 26 | 132 | 5.7 | 7160 | 2 | T27935 | hypothetical prote |
| 27 | 131 | 5.6 | 1257 | 1 | A41060 | neural cell adhesi |
| 28 | 131 | 5.6 | 5175 | 2 | T20992 | hypothetical prote |
| 29 | 131 | 5.6 | 5198 | 2 | T43290 | hemiscentin precure |

| | | | | | | |
|-----|-------|-----|-------|---|--------|---------------------|
| 30 | 130.5 | 5.6 | 304 | 2 | S04663 | T-cell receptor ga |
| 31 | 130 | 5.6 | 1906 | 1 | S68235 | myosin-light-chain |
| 32 | 130 | 5.6 | 4162 | 2 | T42633 | connectin/citin - |
| 33 | 127 | 5.5 | 2629 | 2 | T32735 | telomerase-associ |
| 34 | 127 | 5.5 | 3707 | 2 | S18252 | heparan sulfate pr |
| 35 | 125.5 | 5.4 | 725 | 1 | IJMSNG | neural cell adhesi |
| 36 | 125.5 | 5.4 | 1447 | 2 | A54100 | tumor suppressor p |
| 37 | 125 | 5.4 | 1896 | 2 | T08851 | Down syndrome cell |
| 38 | 125 | 5.4 | 6805 | 2 | S20901 | titin - rabdic (fr |
| 39 | 124 | 5.3 | 1115 | 1 | IJMSNL | neural cell adhesi |
| 40 | 124 | 5.3 | 1298 | 2 | A64157 | hypothetical prote |
| 41 | 124 | 5.3 | 7562 | 2 | I38346 | elastic titin - hu |
| 42 | 123.5 | 5.3 | 323 | 2 | S01895 | T-cell receptor ga |
| 43 | 123.5 | 5.3 | 1091 | 1 | IJCHNL | neural cell adhesi |
| 44 | 122.5 | 5.3 | 4391 | 2 | A38096 | perlecan precursor |
| 45 | 122 | 5.2 | 1323 | 2 | PN0568 | connectin 3B - chl |
| 46 | 121.5 | 5.2 | 338 | 2 | JC4776 | limbic-system-asso |
| 47 | 121 | 5.2 | 946 | 2 | S28061 | SCP1 protein - rat |
| 48 | 121 | 5.2 | 1427 | 2 | I51669 | tumor suppressor - |
| 49 | 120.5 | 5.2 | 333 | 2 | A31923 | amalgam protein pr |
| 50 | 120.5 | 5.2 | 773 | 1 | QRRBG | secretory componen |
| 51 | 120.5 | 5.2 | 3305 | 2 | T18358 | apolipoprotein pr |
| 52 | 119 | 5.1 | 584 | 2 | T08678 | hypothetical prote |
| 53 | 118 | 5.1 | 548 | 2 | JC4917 | signal transducing |
| 54 | 118 | 5.1 | 6642 | 2 | T29757 | protein UNC-89 - C |
| 55 | 117.5 | 5.1 | 858 | 1 | IJRTNC | neural cell adhesi |
| 56 | 117.5 | 5.1 | 1197 | 2 | T30581 | neural cell adhesi |
| 57 | 117.5 | 5.1 | 1232 | 2 | T43027 | neural cell adhesi |
| 58 | 116.5 | 5.0 | 1239 | 1 | A32579 | neuroglial - fruit |
| 59 | 116.5 | 5.0 | 1330 | 2 | S49010 | embryonic receptor |
| 60 | 116 | 5.0 | 458 | 2 | JC1509 | biliary glycoprote |
| 61 | 116 | 5.0 | 521 | 2 | S34338 | neural cell adhesi |
| 62 | 116 | 5.0 | 853 | 1 | IJBONC | neural cell adhesi |
| 63 | 116 | 5.0 | 1021 | 1 | I39207 | leukocyte surface |
| 64 | 115.5 | 5.0 | 1136 | 1 | S57845 | protein-tyrosine k |
| 65 | 115 | 4.9 | 554 | 2 | A69392 | hypothetical prote |
| 66 | 115 | 4.9 | 26926 | 1 | I38344 | titin, cardiac mus |
| 67 | 114 | 4.9 | 257 | 2 | S00682 | IGF Fc receptor al |
| 68 | 114 | 4.9 | 765 | 1 | C42632 | cell adhesion mole |
| 69 | 114 | 4.9 | 812 | 2 | B42632 | cell adhesion mole |
| 70 | 114 | 4.9 | 932 | 2 | A42632 | cell adhesion mole |
| 71 | 113.5 | 4.9 | 245 | 2 | A30154 | IGF receptor alpha |
| 72 | 113.5 | 4.9 | 338 | 2 | JC5519 | 50K glycoprotein p |
| 73 | 113.5 | 4.9 | 1273 | 2 | T42405 | sax-3 protein - Ca |
| 74 | 113.5 | 4.9 | 1277 | 2 | T30532 | neural cell adhesi |
| 75 | 113.5 | 4.9 | 1849 | 2 | C41859 | IGF-specific metal |
| 76 | 113.5 | 4.9 | 2222 | 2 | T13924 | sdk protein - fru |
| 77 | 113 | 4.9 | 540 | 2 | JC4916 | signal transducing |
| 78 | 113 | 4.9 | 729 | 2 | AC3616 | succinoglycan bios |
| 79 | 113 | 4.9 | 1028 | 2 | I58164 | BIG-1 protein - ra |
| 80 | 113 | 4.9 | 1333 | 2 | I78875 | receptor tyrosine |
| 81 | 113 | 4.9 | 2130 | 2 | AB0821 | probable exported |
| 82 | 112 | 4.8 | 588 | 2 | JH0506 | adhesion molecule |
| 83 | 112 | 4.8 | 1091 | 2 | A58532 | glial cell membran |
| 84 | 111.5 | 4.8 | 120 | 2 | S46374 | Ig kappa chain V-J |
| 85 | 111.5 | 4.8 | 422 | 2 | I37891 | interleukin-11 rec |
| 86 | 111.5 | 4.8 | 750 | 2 | S41051 | fibroblast growth |
| 87 | 111 | 4.8 | 340 | 2 | S03517 | T-cell receptor ga |
| 88 | 111 | 4.8 | 514 | 2 | A44100 | cell adhesion mole |
| 89 | 111 | 4.8 | 588 | 2 | A45254 | surface glycoprote |
| 90 | 110.5 | 4.8 | 279 | 2 | C81412 | NOG1/NOG2/bun fam |
| 91 | 110.5 | 4.8 | 1023 | 2 | AB1280 | ATP-dependent dBDN |
| 92 | 109.5 | 4.7 | 584 | 2 | I50419 | 8-glycerin precuro |
| 93 | 109 | 4.7 | 1028 | 2 | A53449 | plasma cytoma-asso |
| 94 | 109 | 4.7 | 1367 | 2 | A41228 | protein-tyrosine k |
| 95 | 108 | 4.6 | 523 | 2 | I50478 | neuroilin - goldfis |
| 96 | 108 | 4.6 | 538 | 2 | I68093 | PRR2 delta - human |
| 97 | 107.5 | 4.6 | 122 | 2 | S40370 | Ig kappa chain - h |
| 98 | 107.5 | 4.6 | 458 | 2 | S68177 | C-CM2a protein is |
| 99 | 107.5 | 4.6 | 519 | 2 | A44786 | ecto-ATPase precu |
| 100 | 107.5 | 4.6 | 1138 | 1 | S24066 | protein-tyrosine k |
| 101 | 107.5 | 4.6 | 1443 | 2 | I50600 | neogenin - chicken |
| 102 | 107.5 | 4.6 | 2629 | 2 | T30987 | telomerase-associ |

```
103 107 4.6 458 1 MMWSR1
104 107 4.6 514 2 A31643
105 107 4.6 521 2 JCI508
106 106.5 4.6 111 2 B37266
107 106.5 4.6 111 2 B37266
108 106.5 4.6 748 2 I38740
109 106 4.6 345 2 S41050
110 106 4.6 345 2 S03199
111 106 4.6 345 2 A46052
112 106 4.6 577 2 I50731
113 106 4.6 1018 2 JCI4211
114 105.5 4.5 293 2 T25836
115 105.5 4.5 238 2 S25741
116 105.5 4.5 598 2 AB1236
117 105 4.5 1150 2 B47114
118 104.5 4.5 345 2 JCI4025
119 104.5 4.5 458 2 S23969
120 104 4.5 515 1 VCIJGA
121 104 4.5 528 2 B75364
122 104 4.5 862 2 I49583
123 104 4.5 1017 2 PC4035
124 104 4.5 1259 2 A43425
125 104 4.5 1268 1 A39640
```

ALIGNMENTS

```
RESULT 1
RMH074
T-cell surface glycoprotein CD4 precursor [validated] - human
N.Alternate names: T-cell surface antigen T4/Leu 3
C.Species: Homo sapiens (man)
C.Date: 28-May-1986 #sequence revision 31-Dec-1988 #text change 20-Apr-2001
C.Accession: A90872; A32722; A34194; A53287; I54176; I54297; A02109; A30039
R.Maddon, P.J.; Littman, D.R.; Godfrey, M.; Maddon, D.E.; Chess, L.; Axel, R.
Cell 42, 93-104, 1985
A.Title: The isolation and nucleotide sequence of a cDNA encoding the T cell surface pro
A.Reference number: A90872; MUID:8554948; PMID:2990730
A.Accession: A90872
A.Molecule type: mRNA
A.Residues: 1-25, 'N', 27-458 <MAD>
A.Experimental source: clone pT48
R.Littman, D.R.; Maddon, P.J.; Axel, R.
Cell 55, 541, 1988
A.Title: Corrected CD4 sequence.
A.Reference number: A90907; MUID:89028665; PMID:3263213
A.Contents: annotation; revision to residue 26
R.Camerini, D.; Seed, B.
Cell 60, 747-754, 1990
A.Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi
A.Reference number: A32722; MUID:90182664; PMID:2107024
A.Accession: A32722
A.Structure: nucleic acid sequence not shown; not compared with conceptual translation
A.Molecule type: mRNA
A.Residues: 26-426, 428-458 <CAM>
R.Carr, S.A.; Hemling, M.E.; Folea-Masserman, G.; Sweet, R.W.; Anumula, K.; Barr, J.R.;
J. Biol. Chem. 264, 21286-21295, 1989
A.Title: Protein and carbohydrate structural analysis of a recombinant soluble CD4 recep
A.Reference number: A34194; MUID:90078232; PMID:2592374
A.Contents: disulfide bonds; carbohydrate-binding sites
A.Accession: A34194
A.Molecule type: protein
A.Residues: 26-394 <CAR>
R.Lederman, S.; DeMartino, J.A.; Daugherty, B.L.; Foeldvari, I.; Yellin, M.J.; Cleary, A
Mol. Immunol. 28, 1171-1181, 1991
A.Title: A single amino acid substitution in a common African allele of the CD4 molecule
A.Reference number: A53287; MUID:92072595; PMID:1961196
A.Accession: A53287
A.Status: not compared with conceptual translation
A.Molecule type: mRNA
A.Residues: 250-264, 'W', 266-280 <LED>
A.Note: sequence extracted from NCBI backbone (NCBIP:68249)
```

```
biliary glycoprote
cell adhesion 80K
biliary glycoprote
Ig kappa chain V r
Ig kappa chain V r
fibroblast growth
opoid-binding pro
vascular cell adhe
Ig heavy chain - n
neural adhesion pr
hypothetical prote
Ig lambda chain -
interleukin protein
phosphoprotein pho
opoid-binding cel
cell-adhesion mole
env polypeptide pr
transcription fact
extracellular solu
differentiation an
cell-cycle-depend
Bravo/Nr-CAM cell
neural cell adhesi
```

```
R.Edwards, M.C.; Gibbs, R.A.
Genomics 14, 590-597, 1992
A.Title: A human dimorphism resulting from loss of an Alu.
A.Reference number: I54176; MUID:93052387; PMID:1330888
A.Accession: I54176
A.Status: translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-72 <RES>
A.Cross-references: GB:U47924; GB:M86525; GB:U72506; NID:G1633547; PIDN:AAB51309.1; PID:
Hum. Immunol. 30, 99-104, 1991
A.Title: Humans with OKT4-epitope deficiency have a single nucleotide base change in the
A.Reference number: I54297; MUID:91216786; PMID:1708753
A.Accession: I54297
A.Status: translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-264, 'W', 266-458 <RE2>
A.Cross-references: GB:M35160; NID:G179143; PIDN:AAA16069.1; PID:G179144
C.Comment: Macrophage tropic strains of HIV-1 bind to a complex of chemokine (C-C) recep
C.Genetics:
A.Gene: GDB:CD4
A.Cross-references: GDB:119767; OMIM:186940
A.Map position: 12pter-12p12
A.Introns: 16/3
A.Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
C.Keywords: AIDS; duplication; glycoprotein; T-cell; transmembrane protein
F.1-25/Domain: signal sequence #status predicted <SIG>
F.26-458/Product: T-cell surface glycoprotein CD4 #status experimental <MAR>
F.134-111/Domain: immunoglobulin homology <IM1>
F.136-166/Domain: immunoglobulin homology #status atypical <IM2>
F.216-239/Domain: immunoglobulin homology <IM3>
F.321-372/Domain: immunoglobulin homology <IM4>
F.397-420/Domain: transmembrane #status predicted <TM>
F.421-458/Domain: intracellular #status predicted <INT>
F.41-109, 135-184, 328-370/Disulfide bonds: #status experimental
F.296, 325/Binding site: carbohydrate (Asn) (covalent) #status experimental
```

```
Query Match 87.4%; Score 2032; DB 1; Length 458;
Best Local Similarity 89.2%; Pred. No. 2, 1e-127;
Matches 405; Conservative 2; Mismatches 17; Indels 30; Gaps 2;
```

```
QY 1 MNRGVPRHLLVQLALPPATQGNKRVLGKGDVLELTCTASQKSIQFMHNSNQIK 60
DB 1 MNRGVPRHLLVQLALPPATQGNKRVLGKGDVLELTCTASQKSIQFMHNSNQIK 60
QY 61 ILNGQSFLLTKGPEKLNDRADSRSLMDQGNFPLILNKLIEDDTYICEVEDQKEVQL 120
DB 61 ILNGQSFLLTKGPEKLNDRADSRSLMDQGNFPLILNKLIEDDTYICEVEDQKEVQL 120
QY 121 LVFGLTNSDTHLQSGSLTTLTSSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTNSDTHLQSGSLTTLTSSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCVLONQKVEFKIDIVLAFQKASSIYKKEGQVERSPFLAFTVEKLTGSGELMW 240
DB 181 TWTCVLONQKVEFKIDIVLAFQKASSIYKKEGQVERSPFLAFTVEKLTGSGELMW 240
QY 241 QAEBAASSKSKSTTDLNKRKESYKRVYQDPRLQNGKLPPLHTLTPOLPOVAGSGNLTLA 300
DB 241 QAEBAASSKSKSTTDLNKRKESYKRVYQDPRLQNGKLPPLHTLTPOLPOVAGSGNLTLA 300
QY 301 LEAKTGKLBQBNLVNWRATQLOKNTLCEVWGPSPKMLSLKLNKEAKSKKEKPYVW 360
DB 301 LEAKTGKLBQBNLVNWRATQLOKNTLCEVWGPSPKMLSLKLNKEAKSKKEKPYVW 360
QY 361 LNPBAGMWQCLSDSGVLTLESNIKVLPTWSTPVHPRASALPAPPTGSLDPOPTASALP 420
DB 361 LNPBAGMWQCLSDSGVLTLESNIKVLPTWSTPVHPRASALPAPPTGSLDPOPTASALP 420
QY 421 DPAASALPALAYISFLIGLGLG-ACVLAATR 453
DB 421 DPAASALPALAYISFLIGLGLG-ACVLAATR 453
QY 397 -----MALIVLGVAAGLPLFGLGIFFCVRCRHR 425
DB 397 -----MALIVLGVAAGLPLFGLGIFFCVRCRHR 425
```

RESULT 2

RMCT4

T-cell surface glycoprotein CD4 - chimpanzee

N/Alternate names: T-cell surface antigen T4/Leu 3

C/Species: Pan troglodytes (chimpanzee)

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999

C/Accession: B32722; A46534

R/Camerini, D.; Seed, B.

Cell 60, 747-754, 1990

A/Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the v

A/Reference number: A32722; MUID:90182664; PMID:2107024

A/Accession: B32722

A/Molecule type: mRNA

A/Residues: 1-432 <CAM>

A/Cross-references: GB:M31135

R/Fomsgaard, A.; Hitech, V.M.; Johnson, P.R.

Eur. J. Immunol. 22, 2973-2981, 1992

A/Title: Cloning and sequences of primate CD4 molecules: diversity of the cellular recep

A/Reference number: A46534; MUID:93049640; PMID:1425921

A/Accession: A46534

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 3-399 <FOM>

A/Note: sequence extracted from NCBI backbone (NCBIP:118332)

C/Comment: This protein is expressed on most thymocytes, as a subset of mature T-cells

C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

C/Keyword: duplication; glycoprotein; T-cell; transmembrane protein

F/-432/Product: T-cell surface glycoprotein CD4 #status predicted <MAT>

F/-371/Domain: extracellular #status predicted <EXT>

F/-86/Domain: immunoglobulin homology <IM1>

F/-111-161/Domain: immunoglobulin homology #status atypical <IM2>

F/-191-274/Domain: immunoglobulin homology <IM3>

F/-396-347/Domain: immunoglobulin homology <IM4>

F/-372-395/Domain: transmembrane #status predicted <TM>

F/-396-432/Domain: intracellular #status predicted <INT>

F/-16-84,130-159,303-345/Diulfide bonds: #status predicted

F/-271,300/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 81.0%; Score 1884; DB 1; Length 432;

Best Local Similarity 87.9%; Pred. No. 1,3e-117;

Matches 346; Conservative 2; Mismatches 20; Indels 30; Gaps 2;

QY 27 KYVLGKKGDTVELTCTASQKKSIOFHWNKSNQIKILGNGSFLTKGPSKLNDRADSRSL 86

DB 2 KYVLGKKGDTVELTCTASQKKSIOFHWNKSNQIKILGNGSFLTKGPSKLNDRADSRSL 61

QY 87 WDOGNFPLIKNLKIEDSDTYICEVEDQKEVQLVFGLTANSDFHLLQGSLITLLESP 146

DB 62 WDOGNFPLIKNLKIEDSDTYICEVEDQKEVQLVFGLTANSDFHLLQGSLITLLESP 121

QY 147 PGSSPSVQCRSPGKNIQGGKTLVSQLELDQSGTCTVLOKQKVEKIDIVLAFOK 206

DB 122 PGSSPSVQCRSPGKNIQGGKTLVSQLELDQSGTCTVLOKQKVEKIDIVLAFOK 181

QY 207 ASSIVYKKEGEVSEFPLAFVTEKLTGSGELMWOAERASSKSWTTPDKNKEVSVKV 266

DB 182 ASSIVYKKEGEVSEFPLAFVTEKLTGSGELMWOAERASSKSWTTPDKNKEVSVKV 241

QY 267 TODPKLQMGKKLPLHLTPALPOYAGSGNLTALBAKTKLHGEVNLVVMRATQLOKL 326

DB 242 TODPKLQMGKKLPLHLTPALPOYAGSGNLTALBAKTKLHGEVNLVVMRATQLOKL 301

QY 327 TCEVWGPTSPKMLSLKLENKAQVSKREKPVAVLNPBAGMOCCLSDSGVLLBSNIV 386

DB 302 TCEVWGPTSPKMLSLKLENKAQVSKREKPVAVLNPBAGMOCCLSDSGVLLBSNIV 361

QY 387 LPTWSTPVHPRASALPAPRTGSALPDPOFASALPDPPASALPALAVISFLIGLIGV- 445

DB 362 LPTWSTPVDP-----MALIVLGVAAGLLFLTGIGIF 392

QY 446 ACVLARTR 453

DB 393 FCVRCRHR 400

RESULT 3

RMWOT4

T-cell surface glycoprotein CD4 - rhesus macaque

N/Alternate names: T-cell surface antigen T4/Leu 3

C/Species: Macaca mulatta (rhesus macaque)

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999

C/Accession: C32722

R/Camerini, D.; Seed, B.

Cell 60, 747-754, 1990

A/Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the v

A/Reference number: A32722; MUID:90182664; PMID:2107024

A/Accession: C32722

A/Molecule type: mRNA

A/Residues: 1-432 <CAM>

A/Cross-references: GB:M31134

C/Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells

C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

C/Keyword: duplication; glycoprotein; T-cell; transmembrane protein

F/-432/Product: T-cell surface glycoprotein CD4 #status predicted <MAT>

F/-371/Domain: extracellular #status predicted <EXT>

F/-86/Domain: immunoglobulin homology <IM1>

F/-111-161/Domain: immunoglobulin homology #status atypical <IM2>

F/-180-293/Domain: immunoglobulin homology <IM3>

F/-396-347/Domain: immunoglobulin homology <IM4>

F/-372-395/Domain: transmembrane #status predicted <TM>

F/-396-432/Domain: intracellular #status predicted <INT>

F/-16-84,130-159,303-345/Diulfide bonds: #status predicted

F/-271,300/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 75.0%; Score 1743; DB 1; Length 432;

Best Local Similarity 81.1%; Pred. No. 2,9e-108;

Matches 347; Conservative 17; Mismatches 34; Indels 30; Gaps 2;

QY 27 KYVLGKKGDTVELTCTASQKKSIOFHWNKSNQIKILGNGSFLTKGPSKLNDRADSRSL 86

DB 2 KYVLGKKGDTVELTCTASQKKSIOFHWNKSNQIKILGNGSFLTKGPSKLNDRADSRSL 61

QY 87 WDOGNFPLIKNLKIEDSDTYICEVEDQKEVQLVFGLTANSDFHLLQGSLITLLESP 146

DB 62 WDOGNFPLIKNLKIEDSDTYICEVEDQKEVQLVFGLTANSDFHLLQGSLITLLESP 121

QY 147 PGSSPSVQCRSPGKNIQGGKTLVSQLELDQSGTCTVLOKQKVEKIDIVLAFOK 206

DB 122 PGSSPSVQCRSPGKNIQGGKTLVSQLELDQSGTCTVLOKQKVEKIDIVLAFOK 181

QY 207 ASSIVYKKEGEVSEFPLAFVTEKLTGSGELMWOAERASSKSWTTPDKNKEVSVKV 266

DB 182 ASSIVYKKEGEVSEFPLAFVTEKLTGSGELMWOAERASSKSWTTPDKNKEVSVKV 241

QY 267 TODPKLQMGKKLPLHLTPALPOYAGSGNLTALBAKTKLHGEVNLVVMRATQLOKL 326

DB 242 TODPKLQMGKKLPLHLTPALPOYAGSGNLTALBAKTKLHGEVNLVVMRATQLOKL 301

QY 327 TCEVWGPTSPKMLSLKLENKAQVSKREKPVAVLNPBAGMOCCLSDSGVLLBSNIV 386

DB 302 TCEVWGPTSPKMLSLKLENKAQVSKREKPVAVLNPBAGMOCCLSDSGVLLBSNIV 361

QY 387 LPTWSTPVHPRASALPAPRTGSALPDPOFASALPDPPASALPALAVISFLIGLIGV- 445

DB 362 LPTWSTPVDP-----MALIVLGVAAGLLFLTGIGIF 392

QY 446 ACVLARTR 453

DB 393 FCVRCRHR 400

RESULT 4

A46254

CD4 precursor - rabbit

C/Species: Oryctolagus cuniculus (domestic rabbit)

CjDate: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
 CjAccession: A46254
 R.Hague, B.F.; Sawesdikosol, S.; Brown, T.J.; Lee, K.; Recker, D.P.; Kindt, T.J.
 Proc. Natl. Acad. Sci. U.S.A. 89, 7963-7967, 1992
 A>Title: CD4 and its role in infection of rabbit cell lines by human immunodeficiency vi
 A:Reference number: A46254; MUID:92390370; PMID:1518821
 A:Accession: A46254
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-459 <HAG>
 A:Cross-references: GB:M57840; NID:9164871; PIDN:AAA11198.1; PID:9164872
 A>Note: sequence extracted from NCBI backbone (NCBIN:112732, NCBI:P:112733)
 CjSuperfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
 F:322-372/Domain: immunoglobulin homology <IMM>

Query Match 49.3%; Score 1147; DB 2; Length 459;
 Best Local Similarity 58.0%; Pred. No. 1,1e-68;
 Matches 240; Conservative 68; Mismatches 94; Indels 12; Gaps 6;

QY 1 MNRGVPFRHL-LIVLQALIPATQGNKVLGKGGDTVELTCTASQKSIQFHWKNSNQ 60
 DB 1 MNRRIYFQCLLVPLALIPATWGTIVRGKAGAVELPCOSSQKRNSEVFMKHAQVK 60
 QY 61 ILGNQGS---SFLTGPSSKLNDRADRSRLMDQGNPPLIKKLTIEDSTYICEVDQKE 116
 DB 61 ILGNQGSSSSFMKGNSEPLSNRVESKKMMQDGSFPLVIXDLRMDSGTYICEVDKXM 120
 QY 117 EVOLLVFGITANSPTHLLOGOSLTLTLESPPGSSPSVQCRSPRGNIQGGKTLSSQEL 176
 DB 121 EVELLVFRITANPNRRLHGGSLTLLESGPSVGSVQKSPENKIIIEGPIKSPKRL 180
 QY 177 QDSGTWCTV-LQNKVEFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGS 235
 DB 181 QDSGTWCHLSFQDQNKELDIKIIVLGPVKASATVYKKEGQVFSPLNFEDSL--S 238
 QY 236 GELMQAERASSSKWITFDLKNKESVYKRVTDPLQWKKLP.LHLTPQALPYAGSG 295
 DB 239 GELMQQVDAASQSWSPSLDRKYSVQKILPDKIQMSKLP.LSTLPQALHRYASG 298
 QY 236 NTLTLEAKTGLHDEVNLVVMRATQLOKNTLCEWGPSPPLTIKLENKEAVSKRE 355
 DB 239 NLTSLTD--KGLHQQVSLVMKLVQVKNKLTCEVLGPDIPMKSLTLEQDEAKVS-TQ 355
 QY 356 KEVWVILPEAGMOCILSDSGVLTLESNIKVLPTWSTVHPASALPAPPTGSA 409
 DB 356 KMWVLDPRAGTWQCLTSGDVLLESKADVLATGSL--HQQPTLLAGLGTGA 407

RESULT 5
 S30193
 T-cell surface glycoprotein CD4 - dog
 CjSpecies: Canis lupus familiaris (dog)
 CjDate: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
 CjAccession: S30193
 R.Milde, K.F.; Conner, G.E.; Mintz, D.H.; Alejandro, R.
 Biochim. Biophys. Acta 1172, 315-318, 1993
 A>Title: Primary structure of the canine CD4 antigen.
 A:Reference number: S30193; MUID:93192324; PMID:7916632
 A:Accession: S30193
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-432 <ML>
 A:Cross-references: EMBL:X6865; NID:9288652; PIDN:CA837664.1; PID:94467377
 CjSuperfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
 F:202-311/Domain: immunoglobulin homology <IMM>

Query Match 47.4%; Score 1101; DB 2; Length 432;
 Best Local Similarity 52.2%; Pred. No. 1,2e-65;
 Matches 235; Conservative 66; Mismatches 97; Indels 52; Gaps 8;

QY 12 LVLLQALIPATQGNKVLGKGGDTVELTCTASQKSIQFHWKNSNQIKILGNQGSFLTK 71
 DB 12 LVLLQALIPATQGNKVLGKGGDTVELTCTASQKSIQFHWKNSNQIKILGNQGSFLTK 71

DB 1 LMLQVLM.PAVT PVREVLVLRAGDAVELPCOTSOKKNIHFNMWRDSMWQILGNQGSFWTV 60
 QY 72 GSKLNDRADRSRLMDQGNPPLIKKLTIEDSTYICEVDQKEVOLLVFGLTAA----- 127
 DB 61 GSKLKRHVESKSLMDQGSFPLVIXDLRMDSGTYICEVDQKEVOLLVFGLTAAKDS 119
 QY 128 -----NSDTHLLOGOSLTLTLESPPGSSPSVQCRSPRGNIQGGKTLSSQELQDSGTW 182
 DB 120 GSSGSSNIRLLOQOQLTTLTLENSGSSPSVQWKGPNKSHGGQNLSTLSPLEQDGTW 179
 QY 183 TCTYLQNKVEFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMQA 242
 DB 180 TCIISQKQVEFNINLVAFQKVSNTFVAREDDQVEFSPLSEDENLV--GELRWQA 237
 QY 243 ERASSKSWITFDLKNKESVYKRVTDPKLQWKKLP.LHLTPQALPYAGSGNLTAL 302
 DB 238 QGASSSLIMISFLENKLSKKEADLPKQKESLP.RFTLPVLSRIASSGLITLNL- 296
 QY 303 AKTGLHDEVNLVVMRATQLOKNTLCEWGPSPDKMLSLKENKEAKVSKREKPVWYN 362
 DB 297 AK-GTIVQEVNLVVMRANSSQNNLTCEVLGPTSPELTSLNLKEQAQKVSQKQLVWVVD 355
 QY 363 PEAGMQCLSDSGVLTLESNIKVLPTWSTVHPASALPAPPTGSALEPDTASALPDP 422
 DB 356 PEGGTWQCLSDKDKVLLASLVN-----SSPV----- 383
 QY 423 PASALPALA-----VISFLGLGLGVAC 447
 DB 384 -VKSWEKFLAITGILGLLILGLCVFC 412

RESULT 6
 A27449
 T-cell surface glycoprotein CD4 precursor - rat
 N:Alternate names: W3/25 antigen
 CjSpecies: Rattus norvegicus (Norway rat)
 CjDate: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Jan-2000
 CjAccession: A27449; A35433
 R.Clark, S.J.; Jeffries, W.A.; Barclay, A.N.; Gagnon, J.; Williams, A.F.
 Proc. Natl. Acad. Sci. U.S.A. 84, 1649-1653, 1987
 A>Title: Peptide and nucleotide sequences of rat CD4 (W3/25) antigen: evidence for deriv
 A:Reference number: A27449; MUID:87175535; PMID:3104900
 A:Accession: A27449
 A:Molecule type: mRNA
 A:Residues: 1-457 <CLA>
 A:Cross-references: GB:M15768; NID:9203387; PIDN:AAA40901.1; PID:9203388
 R.Davis, S.J.; Ward, H.A.; Puklavec, M.J.; Willis, A.C.; Williams, A.F.; Barclay, A.N.
 J. Biol. Chem. 265, 10410-10418, 1990
 A>Title: High level expression in Chinese hamster ovary cells of soluble forms of CD4 T
 A:Reference number: A35433; MUID:90285164; PMID:2113054
 A:Contents: annotation
 CjSuperfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
 CjKeywords: glycoprotein; membrane protein; surface antigen
 F:219-300/Domain: immunoglobulin homology <IMM>

Query Match 43.0%; Score 999; DB 2; Length 457;
 Best Local Similarity 52.9%; Pred. No. 7,4e-59;
 Matches 207; Conservative 61; Mismatches 115; Indels 8; Gaps 5;

QY 1 MNRGVPFRHL-LIVLQALIPATQGNKVLGKGGDTVELTCTASQKSIQFHWKNSNQ 58
 DB 1 MCRGFSFRHLPLLLQLSLKLVVQGTIVLKGEGSABLPCESTSRSSAFMKSSDQ 60
 QY 59 IKILGNQGSFLTGPSSKLNDRADRSRLMDQGNPPLIKKLTIEDSTYICEVDQKEV 118
 DB 61 KTLIGYKKNLIKISLELYSPDRKAWMERGSPPLIINKLRMDSGTYICEVLENKEEV 120
 QY 119 QLVVFGITANSPTHLLOGOSLTLTLES-PPGSSPSVQCRSPRGNIQGGKTLSSQELQ 177
 DB 121 ELWVFRITANPNRRLHGGSLTLTLENSPKVSDPICKKSSIVVDSKAFTHSLRIQ 180
 QY 178 DSGTWCTV-LQNKVEFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGE 237

D8
101 DSGINNCVTYLNQKHSFPMKLTVLGFASITATKAKESESAEFPFLNLTGEESL--OQE 238

Oy
228 LMQAERASSSKSWTFPDLKNREVSVKRATODPKLMGCKPLHLTLTPALPYVAGSNTL 297

D8
229 LRMRKKAPSSQSMTTFFSLKNOKVSVOKSTSNPRQLSFLLPTLIQTPOVISIQFAGSNTL 298

Oy
298 TLALEAKTGKGLHGVNLVMRAITOQN-LTEEWGPFPSPKMLMSLKLENKAIVKSREK 356

D8
299 TLTLD- RGIILVEYNLVMKVIOTPDNSTLTCEWGPFSFKMRLTIKOENOEAFVRROEK 356

Oy
357 PVMVLNPAGMGQCCLSDSQVLLESENIKVL 387

D8
357 VIQVAPEFGAVMOCLISEGEFYKMDSKIQLV 387

RESULT 7

RMSRT4

T-cell surface glycoprotein CD4 precursor - mouse

N|Alternate names: T-cell differentiation antigen U74, T-cell surface antigen T4/Len 3 C|Species: Mus musculus (house mouse)

C|Date: 30-Jun-1987 #sequence revision 30-Jun-1987 #text change 16-Jul-1999 G|Accession: A02110; A26038; A39954; A39955; I54564; I69018; A47642 S|Tourville, B.; Gorman, S.D.; Field, E.H.; Hunkapiller, T.; Parneis, J.R.

A>Title: Isolation and sequence of U74 complementary DNA clones: expression in T cells A|Reference number: A02110; MUID:87018845; PMID:3094146

A|Accession: A02110

A|Molecule type: mRNA

A|Residues: 1-457 <TOU>

A|Cross-references: GB:MJ3816; NID:g192070; PIDN:AAA37267.1; PID:g309112 R|Litman, D.R.; Gettner, S.N.

Nature 325, 453-455, 1987

A>Title: Unusual intron in the immunoglobulin domain of the newly isolated murine CD4 (I) A|Reference number: A26038; MUID:87115821; PMID:3027575

A|Accession: A26038

A|Molecule type: mRNA

A|Residues: 1-457 <LTT>

A|Cross-references: GB:X04836; NID:g50353; PIDN:CMA28539.1; PID:g50354 R|Gorman, S.D.; Tourville, B.; Parneis, J.R.

Proc. Natl. Acad. Sci. U.S.A. 84, 7644-7648, 1987

A>Title: Structure of the mouse gene encoding CD4 and an unusual transcript in brain. A|Reference number: A39893; MUID:88041159; PMID:2823269

A|Accession: A39893

A|Molecule type: DNA

A|Residues: 1-25,'B','27-457 <GR>

A|Cross-references: GB:M17080; GB:J03003; NID:g192515; PIDN:AAA37402.1; PID:g387124 R|Maddon, P.J.; Molinaux, S.M.; Madson, D.E.; Zimmerman, K.A.; Godfrey, M.; Alt, F.W.; Proc. Natl. Acad. Sci. U.S.A. 84, 9155-9159, 1987

A>Title: Structure and expression of the human and mouse T4 genes. A|Reference number: A39955; MUID:88097446; PMID:3501122

A|Accession: A39955

A|Structure: nucleic acid sequence not shown; not compared with conceptual translation

A|Molecule type: mRNA

A|Residues: 25-457 <MD>

A>Note: The cited Genbank accession number, J03564, is not in release .101.0 R|Parneis, J.R.; Hunkapiller, T.

Immunol. Rev. 100, 109-127, 1987

A>Title: U74 and the immunoglobulin gene superfamily: New relationships between the imm A|Reference number: I54564; MUID:86152875; PMID:3328618

A|Accession: I54564

A|Status: translated from GB/EMBL/DDBJ

A|Molecule type: mRNA

A|Residues: 1-457 <RES>

A|Cross-references: GB:M6850; NID:g198670; PIDN:AAA39401.1; PID:g198671 A|Accession: I69018

A>Status: translated from GB/EMBL/DDBU

A|Molecule type: DNA

A|Residues: 208-318 <RD2>

A|Cross-references: GB:M6851; NID:g198672; PIDN:AAA39402.1; PID:g554183 R|Classon, B.J.; Taggartos, J.; Kirszbaum, L.; Maddox, J.; Mackay, C.R.; Brandon, M.; Immunogenetics 23, 129-132, 1986

A>Title: The U74 antigen in mouse and the sheep equivalent are immunoglobulin-like. A|Reference number: AA47642; MUID:86166694; PMID:3082751

A.Accession: A47642
A.Molecule type: protein
A.Residues: 27-43 <CRA>
C.Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells
C.Genetics:
A.Introns: 18/1; 74/1; 128/1; 207/1; 319/1; 386/1; 425/3; 448/2
C.Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
C.Keywords: alternative initiators; duplication; glycoprotein; T-cell; transmembrane protein
F.1-26/Domain: signal sequence #status predicted <Sig>
F.127-457/Product: T-cell surface glycoprotein CD4 #status experimental <MAT>
F.135-114/Domain: immunoglobulin homology <IM1>
F.139-190/Domain: immunoglobulin homology #status atypical <IM2>
F.120-301/Domain: immunoglobulin homology <IM3>
F.1241-457/Product: CD4, brain-specific short form #status predicted <BRA>
F.131-372/Domain: immunoglobulin homology <IM4>
F.139-419/Domain: transmembrane #status predicted <TM>
F.420-457/Domain: intracellular #status predicted <INT>
F.442-112,159-188,328-370/Disulfide bonds: #status predicted
F.187,238,323,392/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 42.7%; Score 993; DB 1; Length 457;
Best Local Similarity 53.8%; Pred. No. 1,9e+59;
Matches 211; Conservative 62; Mismatches 109; Indels 10; Gaps 7;

Dy 1 MNRGVPRH-LLVQLALLPATQGNKVVYLGKKDVELCTASQKSIQFMKNSNQI 59
Db 1 MCRAISLRRLILLQLSLAVTQGKTIVLKGESAEPLPCSSQKIRVFTWKFSDQR 60

Oy 60 KIILNQG-SPLITG--PSKUNDRADSRSLMDQNPFLIKNLKIENSDPYICEVEDQKE 116
Db 61 KIILOHQKGVILIRGSPSQ--DRPDSKKGMEKSGFPLLINKLMESQRYICELERKE 119

Oy 117 EVGLIVEGLTANSTPHLLQGSLLTLPLES--PPSSPFVCGRSPRKNIOGGKITLSVQLE 175
Db 120 EVELMVAEKVTFSPTSLIQOSLTITLDSNSKTSINPLTEKHKKGVASSKYLWSNLR 179

Oy 176 LQDSGTCTVLQNOKKEFEKIDIVLAFFQKASSIVYKKEGOVEBSPFLAFVEXKLTVGS 235
Db 180 VQDSDFPNCTVLTLDQKGMFGMTLSVLGFQSTAITYAKSGESAEEFSPFLNFAEE--NGW 237

Oy 236 GELMWQEPKRSASSSWTTTPDLKRNKSVKRVTOODPKLQMKGKPLPHLTLPOALPOVYGSG 295
Db 238 GELMWKKEKKSFCFPWTFSIFSKKEVSQKSTDEKLQMLKETPLTLTKIPQVSLQFVGSG 297

Oy 296 NLTLTAETAKGTXLQEVNLYVMRATOLQNKLTCEVMGPTSPKMLSLKENKAYSKRE 355
Db 298 NLTITLTD--KGTLHOEYNLYVMKYAQIANLTLTCEWMGPISPQRILTLCQENOZARVSEEQ 355

Oy 356 KPWWVLPNEAGMQCLTSDSGOVLLBSNIKVL 387
Db 356 KVVGVAPEFGTWQCCLISEGDVKYMDSRIQVL 387

RESULT 8
160082
CD4 receptor - human (fragment)
C.Species: Homo sapiens (man)
C.Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 23-Jul-1999
C.Accession: 160082
F.Zverev, V.V.; Sldosov, A.V.; Nedospasov, S.A.; Malushova, V.V.; Udalova, I.A.; Andzhanov, P. Virtual. 40, 100-102, 1995
A.Title: [nucleotide sequence of two exons of the human T-lymphocyte CD4 receptor gene]
A.Reference number: 160082; MUID:95407135; PMID:766667
A.Accession: 160082
A.Status: preliminary; translated from GB/EMBL/DDBB
A.Molecule type: mRNA
A.Residues: 1-71 <RES>
A.Cross-references: GB:S79267; NID:g1086922; PID:AAB35273.1; PID:g1086923
C.Genetics:
A.Introns: 17/1
C.Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

Query Match 15.4%; Score 357; DB 2; Length 71;

Best Local Similarity 98.6%; Pred. No. 2,5e-17;
Matches 70; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMRGVPFRLHLLVLTALPATQGNKVYLGKKGGDTVELTCASSQKKSIOFHMKNQIK 60

Db 1 NMRGVPFRLHLLVLTALPATQGNKVYLGKKGGDTVELTCASSQKKSIOFHMKNQIK 60

QY 61 ILNGSGFLTK 71

Db 61 ILNGSGFLTK 71

RESULT 9

T-cell surface glycoprotein CD4 (allele 1) - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 21-Jan-2000

C:Accession: I47131; S21461

J:Guatafesson, K.; Germana, S.; Sundt, T.M.

J:Immunol. 151, 1365-1370, 1993

A:Title: Extensive allelic polymorphism in the CDR2-like region of the miniature swine C

A:Reference number: I47131; MUID:93329116; PMID:833533

A:Accession: I47131

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-99 <GU2>

A:Cross-references: EMBL:X65629; NID:91928; PIDN:CAA46583.1; PID:9388232

C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

C:Keywords: glycoprotein; T-cell

F/3-81/Domain: immunoglobulin homology <IMM>

Query Match 13.1%; Score 305.5; DB 2; Length 99;

Best Local Similarity 60.2%; Pred. No. 1e-13;

Matches 59; Conservative 16; Mismatches 22; Indels 1; Gaps 1;

QY 32 KKGDTVELTCTASQKKSIOFHMKNQIKILNGSGFL-TKGPSKLNDRADSRSLMDQG 90

Db 1 KAGDLAEPLPCHSSQKKNLPFMKNKSNQTKILGHSFMTASVTELTSLRSLSKKMWHDG 60

QY 91 NFPLIIKNIKEDSDTYICEVEDQKEVQLVFLGTAN 128

Db 61 SFPLIIKNIKEDSDTYICEVEDQKEVQLVFLGTAN 98

RESULT 10

A39016 T-cell surface glycoprotein CD7 precursor - human

M:Alternate names: T-cell leukemia antigen

C:Species: Homo sapiens (man)

C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 05-Nov-1999

C:Accession: A39016; S03520

R:Schuberg, L.E.; Flenoy, D.E.; Kutzberg, J.; Haynes, B.F.; Kaufman, R.E.

Proc. Natl. Acad. Sci. U.S.A. 88, 603-607, 1991

A:Title: Isolation and characterization of the genomic human CD7 gene: structural simila

A:Reference number: A39016; MUID:91110576; PMID:1703303

A:Accession: A39016

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-240 <SCH>

A:Cross-references: GB:M37271; NID:9180163; PIDN:AAA51953.1; PID:9180164

R:Artuffo, A.; Seed, B.

EMBO J. 6, 3313-3316, 1987

A:Title: Molecular cloning of two CD7 (T-cell leukemia antigen) cDNAs by a COS cell expr

A:Reference number: S03520; MUID:88111517; PMID:3501369

A:Accession: S03520

A:Molecule type: mRNA

A:Residues: 1-240 <ARU>

A:Cross-references: EMBL:X06180; NID:929819; PIDN:CAA29546.1; PID:9732757

C:Genetics:

A:Gene: GDB:CD7

A:Cross-references: GDB:119770; OMIM:166820

A:Map position: 17q25.2-17q25.3

A:introns: 28/1

C:Keywords: glycoprotein; membrane protein; surface antigen; tandem repeat

F/1-25/Domain: signal sequence #status predicted <SIG>

F/26-240/Product: T-cell surface glycoprotein CD7 #status predicted <MAT>

F/145-180/Region: 9-residue repeats (X-P-P-X-A-S-A-L-P)

Query Match 12.6%; Score 293; DB 2; Length 240;

Best Local Similarity 40.3%; Pred. No. 2.2e-12;

Matches 91; Conservative 18; Mismatches 49; Indels 68; Gaps 9;

QY 270 PKLQMGKKLPHLLPQALPYAGSNLTLALFAKTHQEVNLYVMARATQKLNLTCE 329

Db 5 PRLTL---LPLLLALAGLPg-----ALAAQEVQSGPHCTTPVGAS---VNITCS 49

QY 330 VMGPTSPKMLSLKFNKAKVSRKRPVWVNLNPEAG----- 366

Db 50 TSGGLRGYLRQL-----GPPQDIIYVEDGVPTTDRFRGRIDFSGSQDNL 98

QY 367 --WMQCLSDSG---QVLESNI-----KVLPT-----W--STPVHPRASALPAPPTG 407

Db 99 ITMRLQLSDGTGTCAITEVNVYSGTLVLTVEQSGQWHRCSDAAPPRASALPAPPTG 158

QY 408 SALPDQPTASALPDPAPASALPALAVISFLGLGACVLAARTR 453

Db 159 SALPDQPTASALPDPAPASALPALAVISFLGLGACVLAARTQ 204

RESULT 11

S21462 T-cell surface glycoprotein CD4 (allele 2) - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 21-Jan-2000

C:Accession: I47132; S21462

J:Guatafesson, K.; Germana, S.; Sundt, T.M.

J:Immunol. 151, 1365-1370, 1993

A:Title: Extensive allelic polymorphism in the CDR2-like region of the miniature swine C

A:Reference number: I47131; MUID:93329116; PMID:833533

A:Accession: I47132

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-99 <GU2>

A:Cross-references: EMBL:X65630; NID:91929; PIDN:CAA46584.1; PID:9388233

C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

C:Keywords: glycoprotein; T-cell

F/3-81/Domain: immunoglobulin homology <IMM>

Query Match 12.1%; Score 280.5; DB 2; Length 99;

Best Local Similarity 56.1%; Pred. No. 4.6e-12;

Matches 55; Conservative 19; Mismatches 23; Indels 1; Gaps 1;

QY 32 KKGDTVELTCTASQKKSIOFHMKNQIKILNGSGFLTKGP-SKLNDRADSRSLMDQG 90

Db 1 KAGDLAEPLPCHSSQKKNLPFMKNKSDQIKILSRHLMHKAASYELSRSLSKKMWHDG 60

QY 91 NFPLIIKNIKEDSDTYICEVEDQKEVQLVFLGTAN 128

Db 61 SFPLIIKNIKEDSDTYICEVEDQKEVQLVFLGTAN 98

RESULT 12

JS0675 vascular cell adhesion molecule-1 precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 05-Nov-1999

C:Accession: JS0675; S19872; S23136

R:Hessien, C.; Moy, P.; Tizard, R.; Chisholm, P.; Williams, C.; Wyse, M.; Burki, L.; Mi

Biochem. Biophys. Res. Commun. 183, 163-169, 1992

A:Title: Cloning of murine and rat vascular cell adhesion molecule-1.

A:Reference number: JS0674; MUID:92181437; PMID:1371918

A:Accession: JS0675

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-739 <HBS>

A:Cross-references: GB:M4488; NID:9207642; PIDN:AAA42332.1; PID:9207643

R. Williams, A. Atkins, R.J. Fries, J. Gimbome, M.A.; Cybulsky, M.I.; Collins, T. submitted to the EMBL Data Library, February 1992

A:Description: Nucleotide sequence of rat vascular cell adhesion molecule-1.

A:Reference number: 519872

A:Accession: S19872

A:Molecule type: mRNA

A:Residues: 1-2, 'G', '4-121, 'HL', '124-165, 'N', '167-738, 'G' <M1>

A:Cross-references: EMBL:X63722; NID:957471; PID:CAA45254.1; PID:957472

R. Williams, A.J. Atkins, R.C.; Fries, J.M.U.; Gimbome Jr., M.A.; Cybulsky, M.I.; Collins, T.

A:Title: Nucleotide sequence of rat vascular cell adhesion molecule-1 cDNA.

A:Reference number: S23136; MUID:92305064; PMID:1377031

A:Accession: S23136

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2, 'G', '4-165, 'N', '167-738, 'G' <M2>

C:Comment: This protein interacts with the beta-1 integrin very late antigen 4 on leukocytes.

C:Genetics:

A:Gene: VCAM-1

C:Superfamily: Immunoglobulin homology

C:Keywords: cell adhesion; transmembrane protein

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-739/Product: vascular cell adhesion molecule 1 #status predicted <VAS>

F:239-293/Domain: immunoglobulin homology <IMM1>

F:328-385/Domain: immunoglobulin homology <IMM2>

F:526-581/Domain: immunoglobulin homology <IMM3>

F:676-696/Domain: transmembrane #status predicted <TRA>

F:697-715/Domain: intracellular #status predicted <INT>

Query Match 7.0%; Score 162; DB 2; Length 739;

Best Local Similarity 22.0%; Pred. No. 0.005;

Matches 86; Conservative 60; Mismatches 155; Indels 90; Gaps 14;

QY 33 KGDVVELTCTASQKSIQFHKNSNQIKILNQGSLFKGPSKLNDRADSRSLMDQGNF 92

DB 238 EGAAYMTQCSSEGLPAEIRFMSK-----LNGVQLL-----SGNA 274

QY 93 PLIKNLIKEDSPTYICE---VEDQKEEVOLLV-----FGLTNSDTHLLQGSLLT 142

DB 275 TLTLIMRMEDSGIYCEGVNLVGRDKTEVELIYQEKFPFVDISPGSQAQVGDSSVLT 334

QY 143 LESPPSSSPVQGRSPGKNIQ-----QKTSVSGLELQDSTWTCTVQLQNKVVF 195

DB 335 CAAVGCDSPFSFWRQTQDSPNGEVRDEGATSTLTLSPVGVEDHSYLCYVCOGRKLEK 394

QY 196 KIDIVLAFQKASSIYKKEGQVEFSFPLA-----FVEXKLTSGS 236

DB 395 TIQVEYYSF-----PEDPEIIEISGLVHGPRVYVNTCTVPNVYPPFDHLEILKERT 445

QY 237 ELWMOAERAS-SSKSWITFDLKNKEVSKRVATODPKLQMGKLT-----PLHLTLPOALP-Q 290

DB 446 TLNKLFLREIGTKS---LETKSLMTFIPYAE-----TGALVCLAKLHSGQMSERPQ 498

QY 291 YAGSGNLTALAEKTKGLQHEVNLVYMRATOLQKLTLEWGPSTSPKMLSLKENKEAK 350

DB 499 ROSTQTLVYVAVAKEPTIIVWSPSPVPEGSPFV---NLTCSSDGFPTPKILWSROLKNGELQ 556

QY 351 VSKREKPVWVLTNPAGWMOCLLSDSGQVILE 381

DB 557 PLUSQ-----NITLSFMAIRKEDSGIYCE 580

RESULT 13

JC2457

vascular cell adhesion protein - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 15-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 08-Oct-1999

C:Accession: JC2457

R.Tsang, Y.T.M.; Haekard, D.O.; Robinson, M.K.

Biochem. Biophys. Res. Commun. 201, 805-812, 1994

A:Title: Cloning and expression kinetics of porcine vascular cell adhesion molecule.

A:Reference number: JC2457; MUID:94271236; PMID:7516159

A:Accession: JC2457

```

A:Molecule type: mRNA
A:Residues: 1-538 <TSA>
A:Cross-references: EMBL: U08351; NID: g9474382; PIDD: AAA21542.1; PID: g9474383
C:Keywords: glycoprotein, transmembrane protein
R:497-517/Domain: transmembrane #status predicted <TMM>
F:75,157,271,330,360/Binding site: carbohydrate (Aan) (covalent) #status predicted

Query Match      6.7%; Score 155.5; DB 2; Length 538;
Best Local Similarity 23.4%; Pred. No. 0.0088;
Matches 65; Conservative 49; Mismatches 91; Indels 73; Gaps 11;

QY 32 KKGPVLLTCTASQKSIQPHWKNINIKLNGSGSLTGKPSKLNLRADSRRLMQGN 91
   ::::: ::::: : : : : : : : : : : : : : : : : : : : : : : :
DB 235 QEGSMWMTCTSEGLPAPQISW-----SKDLNGQQLL--SGN 271
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 92 PFLIKNLIKEDSPYICE---VEDQEEYQVLV-----FGLTANSDHLQGSGLT 141
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 272 ATLILIMRMEDSGIYCEGVNPGVTRKVELVQVAPRDTTISVNPPSTLREGSSVM 331
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 142 TLESPPGSSPSV-----QCRSPRGNKIQGKTLVSQELQDSDGTWCTVLQ---NOKV 193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 332 TCSDGPPAPRIKMSKTLRQGNLEPLSENNTLTLTSMKEDSGIYCEGINQGINRKV 391
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 194 EFKI-----DIVLAFQKAS-----SIYKKE---GEQVEFSPLA 226
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 392 ELIIQAPKDLQTLAFPSSESVKEGDTVIISCTGANNPPTLIILKKKAEKDTVLKSTGDA 451
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 227 FTVEKL-----TGSGELMMQAEARASSSKSWTTPDLKNE 260
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 452 YTHRAFLADGAYECESKNEIGLQTRS-ITLDVKGKE 488
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 14
A41288
vascular cell adhesion molecule 1, long splice form precursor - human
N/Alternate names: VCAM-1
C/Species: Homo sapiens (man)
C/Date: 03-Apr-1992 #sequence revision 03-Apr-1992 #text change 31-Jan-2000
C/Accession: A41288; S11476; A39755; B39755; A61160; A43352; PH1379; A39554
R/Cybulski, M.I.; Fries, J.W.U.; Williams, A.J.; Sultan, P.; Eddy, R.; Byers, M.; Shows
Proc. Natl. Acad. Sci. U.S.A. 88, 7859-7863, 1991
A/Title: Gene structure, chromosomal location, and basis for alternative mRNA splicing c
A/Reference number: A41288; MUID:91352090; PMID:1715583
A/Accession: A41288
A/Molecule type: DNA
A/Residues: 1-739 <CYB>
A/Cross-references: GB:M73255; NID:g340195; PIDN:AAA61270.1; PID:g340196
R/Potter, T.; Newman, W.; Gopal, T.V.
Nucleic Acids Res. 18, 5901, 1990
A/Title: Full length vascular cell adhesion molecule 1 (VCAM-1).
A/Reference number: S11476; MUID:91016951; PMID:1699207
A/Accession: S11476
A/Status: translation not shown
A/Molecule type: mRNA
A/Residues: 1-739 <POL>
A/Cross-references: EMBL:X53051; NID:g37648; PIDN:CAA37218.1; PID:g37649
R/Hession, C.; Tizard, R.; Vassallo, C.; Schiffer, S.B.; Goff, D.; Moy, P.; Chi-Rosso, G
J. Biol. Chem. 266, 6689-6685, 1991
A/Title: Cloning of an alternate form of vascular cell adhesion molecule-1 (VCAM1).
A/Reference number: A39755; MUID:91201302; PMID:1707873
A/Accession: A39755
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 25-646, 648-739 <HES1>
A/Cross-references: GB:M60335
A/Note: The complete translation is not shown
A/Accession: B39755
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-739 <HES2>
A/Cross-references: GB:M60335; NID:g340193; PIDN:AAA61269.1; PID:g340194
A/Experimental source: cell type endothelial cell; tissue type umbilical vein; map 1p32-
R/Cybulski, M.I.; Fries, J.W.U.; Williams, A.J.; Sultan, P.; Davis, V.M.; Gimbrone Jr.,

Am. J. Pathol. 138, 815-820, 1991
 A>Title: Rapid communication. Alternative splicing of human VCAM-1 in activated vascular
 A:Reference number: A61160; MUID:91189297; PMID:1707234
 A:Accession: A61160
 A>Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 25-401, 'T', 403-686 <CY2>
 R:Rademacher, M.F.; McQuillan, J.J.; Rosen, G.D.; Dean, D.C.
 J. Biol. Chem. 267, 16323-16329, 1992
 A>Title: Characterization of the promoter for vascular cell adhesion molecule-1 (VCAM-1)
 A:Reference number: A43352; MUID:92255594; PMID:1379595
 A:Accession: A43352
 A:Molecule type: DNA
 A:Residues: 1-21 <IAD>
 A:Cross-references: GB:M92431; NID:9340197
 A>Note: sequence extracted from NCBI backbone (NCBI:110680, NCBI:110681)
 R:Obborn, L.; Vasallo, C.; Benjamin, C.D.
 J. Exp. Med. 176, 99-107, 1992
 A>Title: Activated endothelium binds lymphocytes through a novel binding site in the alt
 A:Reference number: PH1379; MUID:92308860; PMID:1377228
 A:Accession: PH1379
 A:Molecule type: protein
 A:Residues: 25-181, 'G', 183-402 <OSB>
 C:Comment: This adhesion molecule is induced on endothelial cells by inflammatory cytoki
 C:Genetics:
 A:Gene: GDB:VCAM1
 A:Cross-references: GDB:127922; OMIM:192225
 A:Map position: 1p32-1p31
 C:Keywords: alternative splicing; cell adhesion; glycoprotein; transmembrane protein
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-739/Product: vascular cell adhesion molecule 1, long splice form #status predicted
 F:25-698/Domain: extracellular #status predicted <EXT>
 F:699-720/Domain: transmembrane #status predicted <TM>
 F:721-739/Domain: intracellular #status predicted <INT>
 F:273,365,417,463,531,561/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.7%; Score 155; DB 2; Length 739;
 Best Local Similarity 21.9%; Pred. No. 0.015;
 Matches 91; Conservative 71; Mismatches 166; Indels 80; Gaps 18;

| | | | |
|----|-----|---|-----|
| QY | 8 | RHLVLVQLALPAAT--QGNVVLGKKDYTELTCTASQKSIQFHKNSNQITLGNQ | 65 |
| DB | 211 | RQAVKELQVYISPKQTVISVNPSTKLQEGSVTMTCSSEGLPAPRIFW----- | 258 |
| QY | 66 | GSFLTKPSKLNDRADSRSLMDQGNPLIINKLIEDDTYCE-----VEDQKEVQL | 121 |
| DB | 259 | -----SKLDNGLNQLH---SGNATTLTAMMBSGIYCEGVNLIGNKKEVELI | 307 |
| QY | 122 | V-----FGLTANSDFHLQ--GQSLTTLTSSPPGSSPSVQCRSPRKNIOG-----GKT | 168 |
| DB | 308 | VGKEPFTVAISPGPIAAGIDGVMLTGCVGCESPSPFWMRTQIDSPISGVKRSGETNST | 367 |
| QY | 169 | LSVSQLELDSDGTWCTCTVLQGNKYVEFKIDIVVLAQKASSIVYK---KEGQVEFS--- | 222 |
| DB | 368 | LTLSPVSEFNEHSYCTVYCGHKLEKIQVELYFPRPDELMSGVLNGSSVTSVKV | 427 |
| QY | 223 | ---PPL-AFTVVKLTGSGELTMMQABRASSKSWTF---DLKNKVSVKRTQDPKQ | 273 |
| DB | 428 | BSVYPLDRLEILLGETLL-----ENIEFLEBTDKKSLEKNSLEMTFTPTIED | 476 |
| QY | 274 | MGKCL---PLHLTLPQALPOYAGSGLTLEAKTGKLDHENVLVNRAATOLK---N | 325 |
| DB | 477 | TGKALVCAQAKHIDMEFEPRKQRO---TQTLVYVAVP--RDTYLVSPSSLSIEGSSVN | 531 |
| QY | 326 | LTCBVGWPTSPKLMLSKLEKAEKVKREKPVWVLPNPAWGQCLSDSGQVLLF | 381 |
| DB | 532 | MTCLSGGFPAPKILMSRQLPNEGLOPLSENATLTLISTK-----MEDSGVYLCE | 580 |

C:Species: Homo sapiens (man)
 C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 28-Jan-2000
 C:Accession: S07784; A26883
 R:Barton, C.H.; Dickson, G.; Gower, H.J.; Rowett, L.H.; Plett, W.; Elsom, V.; Moore, S.E.
 Development 104, 165-173, 1988
 A>Title: Complete sequence and in vitro expression of a tissue-specific phosphatidylinos
 A:Reference number: S07784; MUID:89305258; PMID:3253057
 A:Accession: S07784
 A:Molecule type: mRNA
 A:Residues: 1-761 <BAR>
 A:Cross-references: EMBL:X16841; NID:935005; PIDN:CA34739.1; PID:935006
 R:Dickson, G.; Gower, H.J.; Barton, C.H.; Prentice, H.M.; Elsom, V.L.; Moore, S.E.; Cox,
 Cell 50, 1119-1130, 1987
 A>Title: Human muscle neutral cell adhesion molecule (N-CAM): identification of a muscle-
 A:Reference number: A90895; MUID:87301755; PMID:2887295
 A:Accession: A26883
 A:Molecule type: mRNA
 A:Residues: 491-761 <DIC>
 A:Cross-references: GB:M17409; NID:9189097; PIDN:AAA5912.1; PID:9368979
 C:Comment: NCM mediates cell-cell adhesion via homophilic binding with another NCM mol
 C:Genetics:
 A:Gene: GDB:NCAM1; NCAM; CD56
 A:Cross-references: GDB:119448; OMIM:116930
 A:Map position: 11q22.2-11q22.3
 C:Superfamily: neutral cell adhesion molecule; fibronectin type III repeat homology; immu
 C:Keywords: alternative splicing; cell adhesion; duplication; heparin binding; membrane
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-761/Product: neutral cell adhesion molecule phosphatidylinositol-linked form, muscle
 F:34-98/Domain: immunoglobulin homology <IMM1>
 F:113-191/Domain: immunoglobulin homology <IMM2>
 F:152-156/Region: heparin binding #status predicted
 F:161-165/Region: heparin binding #status predicted
 F:228-289/Domain: immunoglobulin homology <IMM3>
 F:263-272/Region: NCM binding #status predicted
 F:322-387/Domain: immunoglobulin homology <IMM4>
 F:419-481/Domain: immunoglobulin homology <IMM5>
 F:499-587/Domain: fibronectin type III repeat homology #status atypical <FN3A>
 F:633-720/Domain: fibronectin type III repeat homology #status atypical <FN3B>
 F:41-96, 139-189, 235-287, 329-385, 426-479/Diulfide bonds: #status predicted
 F:222,315,347,423,449,478/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.2%; Score 145; DB 1; Length 761;
 Best Local Similarity 20.5%; Pred. No. 0.07;
 Matches 100; Conservative 67; Mismatches 211; Indels 110; Gaps 20;

| | | | |
|----|-----|---|-----|
| QY | 3 | RG-VPRHLLVLQALPAATQGNKVLGKKDYTELTCTASQKSIQFHKNSNQITKI | 61 |
| DB | 196 | RGEINFDIOIVTVVPTIQARQNI VNAATNLGQSVTLVCDAGEPPEPTSMW----- | 247 |
| QY | 62 | LGNGSFLTKPSKLNDRADSRSLMDQGNPLIINKLIEDSDTYICEVE-----DQKEE | 117 |
| DB | 248 | -----TKDGEIIBQEBDEDEKIFSDSSQLTKKYDKKDAEYICIAENKAGEQDAT | 299 |
| QY | 118 | VQLLVFG---LTRANSDFHLQ--GQSLTTLTSSPPGSSPSVQCR-----SPRKNIOG- | 166 |
| DB | 300 | IHLVFAKPKITVYENQTAEMLEBOVTLTCEASDPISTIWRSTRNISSEBKTLDQHM | 359 |
| QY | 167 | -----KTLVSQLELDSDGTWCTCTVLQ--NQKVEFKIDIVVLAQKASSIVYKKEG | 216 |
| DB | 360 | VVRSHARVSLTLLKSISYDTAGETICTASNTIGDSQSMYLEVOYAPRLQGPVAVYTWEG | 419 |
| QY | 217 | EQVE-----PSPLAFTVVKLTGSGELTMMQABRASSKSWTFPDKN-KEVSVVRVTD | 270 |
| DB | 420 | NOVAITECVFAYPSA-TIS-----WFRDQGLLPSNSNINIKIYNTPSASYLEVTPDS | 470 |
| QY | 271 | KLQNGK-----KLPLHLTLPQALPOYAGSGLTLEAKTGKLDHENVLVNRAATOLK | 310 |
| DB | 471 | ENDFGNNCTAVNRIGQESLEFLVQADTPSSPSIDQVEPVSSTAQVQFDEPBAATG---- | 526 |
| QY | 311 | ENVLVNRAATOLQKLTCEVWGPTSPKLMLSKLEK- EAKVSKREKPVWV--LNPBAG | 366 |
| DB | 527 | --GVPILKYKAEMRAVGEVW-----HSRWYDAKESMGEIVITVIGLKPEPTT | 571 |

Db 719 DVGSGNETNNVITWKLRLMMDNAPQIQXRVQKPLGQETWKEQTVSDPFLVNSNTS 778
QY 250 SWTFEDLNKKEYSVKRVYQDDPKLQNG 275
Db 779 TFVPEIKVQAVNQGKGPFPQVITG 804
RESULT 21
S05479
neural cell adhesion molecule L1 precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: S05479; B60850; S22167
R/Moore, M.; Tacke, R.; Scherer, H.; Teplow, D.; Frueh, K.; Schachner, M.
Nature 334, 701-703, 1988
A/Title: Neural adhesion molecule L1 as a member of the immunoglobulin superfamily with
A/Reference number: S05479; MUID:88318924; PMID:3412448
A/Accession: S05479
A/Molecule type: mRNA
A/Residues: 1-1260 <MOO>
A/Cross-references: EMBL:X12875; NID:953336; PID:CAA33368.1; PID:953337
A/Note: the authors translated the codon CCT for residue 166 as leu, ACT for residue 396
A/Note: part of this sequence, including the amino end of the mature protein, was confir
R/Rathjen, F.G.; Wolff, J.M.; Frank, R.; Bonhoeffer, F.; Rutishauser, U.
J. Cell Biol. 104, 343-353, 1987
A/Title: Membrane glycoproteins involved in neurite fasciculation.
A/Reference number: A60850; MUID:87109457; PMID:3805123
A/Accession: B60850
A/Molecule type: protein
A/Residues: 20-28,'XX','31-36 <RAT>
R/Kohl, A.; Gleese, K.P.; Mohajer, M.H.; Montag, D.; Moore, M.; Schachner, M.
submitted to the EMBL Data Library, December 1991
A/Description: Analysis of promoter activity and 5' genomic structure of the neural cell
A/Reference number: S22167
A/Accession: S22167
A/Molecule type: DNA
A/Residues: 1-165,'L',167-189,'E',191-281,'S',283-395,'S',397-514,'ABEKNPVDV',524,'GEGNE
A/Cross-references: EMBL:X63511
C/Genetics:
A/Introns: 26/1, 31/1, 66/2, 133/1, 174/1, 231/1, 268/2, 330/1, 374/1, 422/1, 459/2
A/Note: the list of introns may be incomplete
C/Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology;
C/Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; transmembran
F.1-19/Domain: signal sequence #status predicted <SIG>
F.20-1260/Product: neural cell adhesion molecule #status experimental <MAT>
F.256-313/Domain: immunoglobulin homology <IMM1>
F.440-498/Domain: immunoglobulin homology <IMM2>
F.531-592/Domain: immunoglobulin homology <IMM3>
Query Match 5.8%; Score 134.5; DB 1; Length 1260;
Best Local Similarity 20.1%; Pred. No. 0.69;
Matches 73; Conservative 59; Mismatches 140; Indels 91; Gaps 15;
QY 34 GDTVELTCTAGSKSIQFHWKNSNOIKILNGSGFLTGKPSKLNDRADRSRLMDQGNP 93
Db 346 GETALDLCQVQRPQPEITWR-----INGMSWETVKNQDKYRI-BQGS-- 387
QY 94 LIKKLKIEDSDTYICEVEDOK-----EENGLVFGLTANSDTHL-DQGSGLTTL 143
Db 388 LILSNVQPTDWTWTCCEARNQGLLIANNAYIVVLPARILTKDKQTMAAEGSTAYILC 447
QY 144 ESPSPSPVQCRSPRKNI-----QGGKTLVSQLEDSGSGTWCTVQLQNKVEF 195
Db 448 KAFGAPVPVQWLDEGTVLQDERFPYANGTILSRDLQANDTGRYCOQANDQNTI 507
QY 196 KIDIVLAFQKASSI-----VYKKEGEVEFSPLAFTVEKLTGSGELMWAQERASSSK 249
Db 508 LANLVQ---KEATQITQGPBSAIEKKGARVTFQASFPSP---QASITWGDGR----- 557
QY 250 SWITFDLNKKEYSVKRVYQDDPKLQNGKULPLHLTPQALPQVAGSGNLTALAEATGKLH 309
Db 558 -----DLQERGDSDRYFIEDGLVIT-----QSL-DYSDQGNYSVASTELDEVE 600

QY 310 QEVLVYWMRA-----TQLQKNTLCEWGP-----TSPKLMLSLKLENKEAK 350
Db 601 SBAQLLVGSPGPVPHLELSDRHLLKQSQVHLS---WSPADHNSPIEKYDIEFEDKEMA 657
QY 351 VSK 353
Db 658 PEK 660
RESULT 22
T34416
hypothetical protein F12F3.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence revision 29-Oct-1999 #text_change 29-Oct-1999
C/Accession: T34416
R/Fulton, B.; Mohlmann, P.
submitted to the EMBL Data Library, July 1998
A/Description: The sequence of C. elegans cosmid F12F3.
A/Reference number: 221521
A/Accession: T34416
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2783 <FUL>
A/Cross-references: EMBL:U80022; PID:AC25886.1; GSPDB:GN00023; CESP:F12F3.2
A/Experimental source: strain Bristol N2; clone F12F3
C/Genetics:
A/Gene: CESP:F12F3.2
A/Map position: 5
A/Introns: 45/3; 90/3; 451/3; 509/1; 2313/3; 2341/3; 2378/3; 2414/2; 2453/3; 2474/2; 252
Query Match 5.7%; Score 133; DB 2; Length 2783;
Best Local Similarity 20.7%; Pred. No. 2.5;
Matches 94; Conservative 68; Mismatches 161; Indels 132; Gaps 22;
QY 28 VVLGKKDPTVELTCTAGSKSIQFHWKNSNOIKILNGSGFLTGKPSKLNDRADRSRLW 87
Db 1187 VVLTAETATATFTQSYANPAQVW-----LHNGKALQQTNSYKTRLF 1231
QY 88 DQGNPFLIKKLIKIEDSDTYICEVEDQKEV-----QLLVFGLTANS----- 129
Db 1232 DNTATLVIEVWDELGGTTAAVANNOPGDVHTSAQLTISGEAKKTAASLPYITIEKP 1291
QY 130 DTHLLQSQSLTLLTLESPGSSP-----SVQCRSP-----GKNIQSGKTLVSQ 173
Db 1292 KINVEG--ATLSIQADLNSGPPEVWMLKDNSELVSDRQKCDGVNYG---LVVRD 1345
QY 174 LELDQSGTWCTVQLQNKVEFKIDIVLAFQKASSIVYKKEGEVEF-----S 222
Db 1346 VGLDEBGTYYITTAENEGKIRQNTVEVS---TKSEYVEKKEKKVEKQGGKKKPRFG 1402
QY 223 FPL---AFTVEKLT-----GSGELMWAQERASSSKSWITF-DLNKKEYSVK----- 264
Db 1403 LPRPSGASKTEQVTMAFDAPSEGPADSYEVRRCPPDQREWVSCGTSLSLELITGLPNT 1462
QY 265 ---RVTDPPQLQNGKULPLHLTLPCA---LPOYASGNLTALAEATGKLHOBVNLV 316
Db 1463 EYIFRVAQKKNQGGSEMSMTLKTASVGAQPF-----TISPGSK-----IIA 1507
QY 317 MRATQLOKNTLCEWGPSPKLMLSLKLENKEAVYSREK-----PWTYLN---PEA 365
Db 1508 NRDEFE--IAVEFSGPTP---SVKMYKENLQIVDEKIDAVTSTSSILNKSQEN 1561
QY 366 GMMQCLL-SDSGOVLLESNI---KVLPTWSTPVH 395
Db 1562 GTFNCLIBNELGQASACQVITFNKPSLQSTPDH 1596
RESULT 23
T13669
neuromusculin - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 13-Aug-1999 #sequence revision 13-Aug-1999 #text_change 17-Nov-2000
C/Accession: T13669

R.Kanta, A.; Han, P. L.; Kim, Y. T.; Bellen, H.
Neuron 11, 673-687, 1993
A:Title: Neuromusculin, a Drosophila gene expressed in peripheral neuronal precursors and in the CNS
A:Reference number: Z17697; MUID:94000831; PMID:8398154
A:Accession: T13669
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1011 <KAN>
A:Cross-references: EMBL:L23146; NID:g385073; PID:g385074; PIDN:AAA03750.1
C:Genetics:
A:Gene: nrm
A:Cross-references: FlyBase:FBgn0005629

```

Query Match 5.7%; Score 132; DB 2; Length 1011;
Best Local Similarity 21.1%; Pred. No. 0.75;
Matches 83; Conservative 68; Mismatches 141; Indels 102; Gaps 20;

OY 10 LLLVLQALLPAAT-OGNKVVLGKKGDVETCTASQKKSIOFH----WKNNOI-KIL 62
      ||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 26 LVLVLCTLAVDSSTQAVDTTISQSGSVLPFCPVDAEKCKGLSLMFLFKGDDRIIAAML 85
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 63 GNQGSFLTKGSGKNDRADSRSLMDGNFLLIKNKLEBSPIYICEVEBQKEEVQLLV 122
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 86 GD-----SNVTSVAKFERVATV-EQNPYRLVTKDLKIAEDVLYLDT-----T 128
      : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 123 FGLTANSDTHLQGSILFTLESPPGSSPSVQCRSPRKNIQGGKTLISVSLFODSGTW 182
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 129 FPIPETCDN-FNGRIELRLVLP--TEVILIDAKDRIRKNSV--VGMQGRQSLKA 182
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 183 TCTYLQNOKKVFEKIDIVLAFQKASSIVYKKEGBQVEFSPFLAFYEKL-TSGGELMWQ 241
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 TCTYANTPRQEB-----VSMFRGKRLITTVSPTHDLVDGLYTSLEIEDWT 227
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 242 AERASSRSKWITFDLKNKSVSKRVTO--DPKLOM-----GKKLPLHL 282
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 228 LSRRELMAD--IECRVSAAIQWTVTKFSVDLVQVPTSIDINGVKHHYQGSKVLTIC 284
      : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 283 TLPAQLPQY-----AGSGNLT---LALFAKTKLHQEVLVLMVRAATQLQKN 325
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 285 DIHGARPANVLTWYVTTTIISSGENEIRTEVRSKSLSEKSDGTFHQSLEL-FPNATRFENDR 343
      : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 326 -LTCEVMGPTSPKMLSLKLENKAKVSKREKPV 358
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 344 VFRCEA-----ENIVLQIN-REKPI 362
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 24
A:Accession: AB8852
Protein unc-22 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: AB8852
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:9905613; PMID:9851916
A:Note: See websites genome.wustl.edu/genC/elegans/ and www.sanger.ac.uk/Projects/C.elegans/
A:Note: published errata appeared in Science 283, 35, 1999; and
A:Accession: AB8852
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-6831 <STO>
A:Cross-references: GB:chr_IV; PIDN:CAA98081.1; PID:g3881830; GSPDB:GN00022
C:Genetics:
A:Gene: unc-22
A:Map position: 4
A:Superfamily: twt1ch1n; fibronectin type III repeat homology; immunoglobulin homology;

```

Query Match 5.7%; Score 132; DB 2; Length 6831;
Beet Local Similarity 21.8%; Pred. No. 9.8;
Matches 97; Conservative 66; Mismatches 168; Indels 114; Gaps 23;

```

Db      986 RHITVV-----PKSQ-KV---NESDLATLETVDNDKDAEIVVMHDKRIDIDGVK-- 1032
Qy      68 FLTGPSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEQDKEVEQLVFG--- 124
Db      1033 -----FKVESNNKRR-----LIINGARIEDGEVKCTTKDKDRTAQQLIVDAKKN 1077
Qy      125 -LTANSDTHLLQOGSLTFLTESPPGSSPVQCSPRKNIQ---GSK-----TLS 170
Db      1078 FIVALKDTEVIEKDDVILMCQTKDTKTPGIMFRN--SKQISMBGKFKETQSRNGHTHLK 1135
Qy      171 VSQLELQDSGWITCVLQNGKKVEKIDIVLAFQKASSIVYK-----KEGEVEESFP 224
Db      1136 IGKIEKMEADYEL---DQAGLRGSCNVYLEAEKRPILNWKPKKIEAAGBEPVAVKP 1191
Qy      225 LAFVTEKLTSGSELIMQOAEERASSSEKSWITFDLKNKEVSVKRVQDPRKLOMGKPLPHLTL 284
Db      1192 FQI---KQTRGD-----PKAQIILKNGKPRIDEBEKKL-VEVIIKDDVAELVFK----- 1235
Qy      285 PQALPQVAGSGNLTALAEAKTKLHOEVNLVVM-----RATOLOKNLTCE---VMGPT 334
Db      1236 ---NPQLADNGKMALELGNSAGTALAPFELFVDDKPKRPGRPLETKNVTBEGDLVWGTP 1292
Qy      335 SPKMLSLK---LENKE-----AKVSK-----REKPVVLNPEAGMMOQLSD 374
Db      1293 DPDEGAVPVKAVIIMQEGRGSGNMAKVGETKGTDFKVDLKEHGGEKFRVAK-INECGLSD 1351
Qy      375 --SGQVLTESNIKVLPTWSTPVIHPR 397
Db      1352 PLTESVLAKN---DYGVGKPK 1371

```

RESULT 25

S57242

twichin [similarity] - Caenorhabditis elegans

N/Alternate names: myosin-regulating protein

N/Contains: protein kinase (EC 2.7.1.-)

C/Species: Caenorhabditis elegans

C/Date: 28-Oct-1995 #sequence revision 24-Oct-1997 #text change 20-Jun-2000

C/Accession: S57242, S07571, S06797, S57218, T27933, T28050

R/Bentall, G.M.; L'Hernault, S.W.; Morris, M.E.

R/Submitted to the EMBL Data Library, February 1993

A/Description: Additional sequence complexity within twitching of Caenorhabditis elegans

A/Reference number: S57242

A/Accession: S57242

A/Molecule type: DNA

A/Residues: 1-683 <BEN1>

A/Cross-references: EMBL:L10351

A/Experimental source: var. Bristol

R/Bentall, G.

R/Submitted to the EMBL Data Library, November 1989

A/Reference number: S07571

A/Accession: S07571

A/Molecule type: DNA

A/Residues: 792-6839 <BEN2>

A/Cross-references: EMBL:X15423; NID:g6897; PIDN:CA33463.1; PID:g6898

A/Experimental source: var. Bristol

R/Bentall, G.M.; Kiff, J.E.; Neckelmann, N.; Moerman, D.G.; Waterston, R.H.

Nature 342, 45-50, 1989

A/Title: Sequence of an unusually large protein implicated in regulation of myosin activation

A/Reference number: S06797; MUID:90044042; PMID:2812002

A/Accession: S06797

A/Status: nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 806-1175,1178-1998, 'Y', 2000-3040, 'I', 3042-3335, 'I', 3337-5693,5696-6359, 'I', 6361-6839

A/Cross-references: EMBL:X15423

A/Experimental source: var. Bristol

R/Bentall, G.M.; L'Hernault, S.W.; Morris, M.E.

Genetics 134, 1097-1104, 1993

A/Title: Additional sequence complexity in the muscle gene, unc-22, and its encoded protein

A/Reference number: S57218; MUID:93387664; PMID:8397135

A/Accession: S57218

A/Molecule type: DNA

A:Residues: 2-99;108-194,'Q',196-206;374-468;658-753 <BEN4>
 A:Experimental source: var. Bristol
 R:White, S.
 Submitted to the EMBL Data Library, May 1996
 A:Reference number: Z20442
 A:Accession: T27934
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 'MGIPGKKCKQ',19-6839 <W12>
 A:Cross-references: EMBL:Z73897; PIDN:CAA98064.1; GSPDB:GN00022; CESP:ZK617.1a
 A:Experimental source: clone ZK617
 R:Harris, B.
 Submitted to the EMBL Data Library, May 1996
 A:Reference number: Z20458
 A:Accession: T28030
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 'MGIPGKKCKQ',19-6839 <W12>
 A:Cross-references: EMBL:Z73899; PIDN:CAA98081.1; GSPDB:GN00022; CESP:ZK617.1a
 A:Experimental source: clone ZK829
 C:Comment: Lack of unc-22 leads to a constant twitching of the body muscles.
 C:Genetics:
 A:Gene: unc-22; CESP:ZK617.1a
 A:Map position: 4
 A:introns: 18/3; 69/3; 143/2; 176/3; 264/2; 387/3; 413/2; 471/1; 516/3; 550/3; 582/3; 60152/3; 6691/3; 6776/1; 6808/3
 C:Superfamily: twitchin; fibronectin type III repeat homology; immunoglobulin homology;
 C:Keywords: ATP, autophosphorylation, duplication, muscle, phosphotransferase, serine/th
 F:806-898,899-990,991-1083,1084-1175,1178-1273,1474-1567,1770-1864,2066-2158,2358-2450,2
 96-5750,6263-6356,6386-6478,6541-6635,6649-6742,6745-6838/Region: motif 2
 F:1274-1372,1373-1473,1568-1670,1671-1769,1865-1964,1965-2065,2159-2258,2259-2357,2451-2
 23,4215-4313,4314-4415,4416-4516,4612-4710,4711-4811,4908-5009,5010-5109,5110-5210,5399-
 F:5940-6197/Domain: protein kinase homology <KIN>
 F:5940-5956/Region: protein kinase ATP-binding motif
 F:5971/Active site: Lys #acatus predicted

Query Match 5.7%; Score 132; DB 2; Length 6839;
 Best Local Similarity 21.8%; Pred. No. 9.8;
 Matches 97; Conservative 66; Mismatches 168; Indels 114; Gaps 23;

QY 8 RHLLVLTQLALLPATQGNKVVLLGKKGVLTCTASQKSIQFHWKNSNOIKILGNQGS 67
 DB 994 RHTFVV-----PMKSC--KV---NESDLATLETVDNDKDAEVMWMDKRIDIDGVK-- 1040
 QY 68 FLTKGPKSLNDRADSRSLMDQGNPFLIKLIKLEDSPTYICEVDQKEVOLLVFG--- 124
 DB 1041 -----FKVSSNRKR-----LINGARIEDHGEYCTTKDDRTMAQLIVDAK 1085
 QY 125 -LTNSDTHLQGSILTLTLESPGSSPSVOCSPRGKNIQ---GSK-----TIS 170
 DB 1086 FIVALKDTEVIEKDVTLMCQTKDTGTFGRN--GKQISSMPGSKFETQSRNGHTLTK 1143
 QY 171 VSQLELDQSGTCTVVLQNKVEFKIDIVLAFQKASSIVYK-----KEGEVFSFP 224
 DB 1144 IKIENENADYVEI---DQAGLRGSCNVTLEAEKRPILNMKPKKIEAKGEPGVKVP 1199
 QY 225 LAFVTEKLTGSGELMWQAEARASSSKSWITPDLKKNKESVYKVTQDPKLOMKKLPILHLTL 284
 DB 1200 FOI---KSTRRGD-----PKQILKNGKPIDEMRKLVVEYITKDDVAEIVFK----- 1243
 QY 285 POALPOVAGSNLTALAEAKTKLHOEVNLVVM-----RATOLQKNLTCE---VMGPT 334
 DB 1244 ---NPQLADTGKMALELGNAGTALAPFLFKDKRPKPKPLETKNTVTAEGLDLVWGTP 1300
 QY 335 SPKMLSLK---LENKE-----AKYSK-----REKPVWVNLPEAGMOCCLSD 374
 DB 1301 DPDEGAPVAVYIIEHQEGRSQWMAKVGETKTDFFVKDLKEHGEYKFRVKA-LNECGLSD 1359
 QY 375 --SGQVLESNIKVLPTWSTPVHPR 397
 DB 1360 PLTGESVLAKN-----PYGVGPKPK 1379

RESULT 26
 T27935
 hypothetical protein ZK617.1b - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2000
 C:Accession: T27935; T28031
 R:White, S.
 Submitted to the EMBL Data Library, May 1996
 A:Reference number: Z20442
 A:Accession: T27935
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-7160 <W12>
 A:Cross-references: EMBL:Z73897; PIDN:CAA98065.1; GSPDB:GN00022; CESP:ZK617.1b
 A:Experimental source: clone ZK617
 R:Harris, B.
 Submitted to the EMBL Data Library, May 1996
 A:Reference number: Z20458
 A:Accession: T28031
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-7160 <W12>
 A:Cross-references: EMBL:Z73899; PIDN:CAA98082.1; GSPDB:GN00022; CESP:ZK617.1b
 A:Experimental source: clone ZK829
 C:Genetics:
 A:Gene: CESP:ZK617.1b
 A:Map position: 4
 A:introns: 10/3; 61/3; 135/2; 168/3; 256/2; 379/3; 405/2; 463/1; 508/3; 542/3; 574/3; 593067/1; 3141/3; 3269/1; 6473/3; 7012/3; 7097/1; 7129/3
 C:Superfamily: twitchin; fibronectin type III repeat homology; immunoglobulin homology;

Query Match 5.7%; Score 132; DB 2; Length 7160;
 Best Local Similarity 21.8%; Pred. No. 10;
 Matches 97; Conservative 66; Mismatches 168; Indels 114; Gaps 23;

QY 8 RHLLVLTQLALLPATQGNKVVLLGKKGVLTCTASQKSIQFHWKNSNOIKILGNQGS 67
 DB 1315 RHTFVV-----PMKSC--KV---NESDLATLETVDNDKDAEVMWMDKRIDIDGVK-- 1361
 QY 68 FLTKGPKSLNDRADSRSLMDQGNPFLIKLIKLEDSPTYICEVDQKEVOLLVFG--- 124
 DB 1362 -----FKVSSNRKR-----LINGARIEDHGEYCTTKDDRTMAQLIVDAK 1406
 QY 125 -LTNSDTHLQGSILTLTLESPGSSPSVOCSPRGKNIQ---GSK-----TIS 170
 DB 1407 FIVALKDTEVIEKDVTLMCQTKDTGTFGRN--GKQISSMPGSKFETQSRNGHTLTK 1464
 QY 171 VSQLELDQSGTCTVVLQNKVEFKIDIVLAFQKASSIVYK-----KEGEVFSFP 224
 DB 1465 IKIENENADYVEI---DQAGLRGSCNVTLEAEKRPILNMKPKKIEAKGEPGVKVP 1520
 QY 225 LAFVTEKLTGSGELMWQAEARASSSKSWITPDLKKNKESVYKVTQDPKLOMKKLPILHLTL 284
 DB 1521 FOI---KSTRRGD-----PKQILKNGKPIDEMRKLVVEYITKDDVAEIVFK----- 1564
 QY 285 POALPOVAGSNLTALAEAKTKLHOEVNLVVM-----RATOLQKNLTCE---VMGPT 334
 DB 1565 ---NPQLADTGKMALELGNAGTALAPFLFKDKRPKPKPLETKNTVTAEGLDLVWGTP 1621
 QY 335 SPKMLSLK---LENKE-----AKYSK-----REKPVWVNLPEAGMOCCLSD 374
 DB 1622 DPDEGAPVAVYIIEHQEGRSQWMAKVGETKTDFFVKDLKEHGEYKFRVKA-LNECGLSD 1680
 QY 375 --SGQVLESNIKVLPTWSTPVHPR 397
 DB 1681 PLTGESVLAKN-----PYGVGPKPK 1700

RESULT 27
 A41060
 neural cell adhesion molecule L1 precursor - human
 N:Alternate names: L1CAM
 C:Species: Homo sapiens (man)

C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #ext change 21-Jul-2000
 C/Accession: A41060; S18454; A35331; S21971; S21972; A60223; A31072; G02506
 R/Havlin, M.L.; Lemmon, V.
 Genomic 11, 416-423, 1991
 A/Title: Molecular structure and functional testing of human L1CAM: an interspecies comp
 A/Reference number: A41060; MUID:9210663; PMID:1169655
 A/Accession: A41060
 A/Molecule type: mRNA
 A/Residues: 1-1257 <HLA>
 A/Cross-references: GB:M64296; NID:g186053; PIDN:AA014352.1; PID:g3068548
 R/Kobayashi, M.; Minura, M.; Asou, H.; Oyemura, K.
 Biochim. Biophys. Acta 1090, 238-240, 1991
 A/Title: Molecular cloning of cell adhesion molecule L1 from human nervous tissue: a com
 A/Reference number: S18454; MUID:9203169; PMID:1932117
 A/Accession: S18454
 A/Molecule type: mRNA
 A/Residues: 1-3, 'V', '5-215, 'T', '217-249, 'T', '251-275, 'SV', '278-356, 'E', '358-625, 'V', '627-1257
 A/Cross-references: EMBL:X59847; NID:g35009; PIDN:CAA42508.1; PID:g35010
 A/Note: the authors translated the codon GAA for residue 27 as GLY
 R/Djabali, M.; Mettel, M.G.; Nguyen, C.; Roux, D.; Demengeot, J.; Denizot, F.; Moos, M.;
 Genomics 7, 587-593, 1990
 A/Title: The gene encoding L1, a neural adhesion molecule of the immunoglobulin family,
 A/Reference number: A35331; MUID:90353957; PMID:2387585
 A/Accession: A35331
 A/Molecule type: DNA
 A/Residues: 332-371 <DNA>
 A/Cross-references: GB:M55271
 R/Rosenthal, A.; Mackinnon, R.N.; Jones, D.S.C.
 Nucleic Acids Res. 19, 5395-5401, 1991
 A/Title: PCR walking from microdissection clone M54 identifies three exons from the huma
 A/Reference number: S21971; MUID:9202023; PMID:1933824
 A/Accession: S21971
 A/Molecule type: DNA
 A/Residues: 1082-1176 <ROS>
 A/Cross-references: EMBL:X58775; NID:g29642; PIDN:CAA41576.1; PID:g29643
 A/Accession: S21972
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: mRNA
 A/Residues: 353-935, 'V', '937-1176 <ROS>
 A/Cross-references: EMBL:X58776; NID:g29644; PIDN:CA837831.1; PID:g4467833
 R/Harper, J.R.; Prince, J.T.; Healy, P.A.; Stuart, J.K.; Nauman, S.J.; Stallcup, W.B.
 J. Neurochem. 56, 797-804, 1991
 A/Title: Isolation and sequence of partial cDNA clones of human L1: homology of human ar
 A/Reference number: A60223; MUID:9132183; PMID:1993855
 A/Accession: A60223
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1030-1115, 'WLC', '1118-1176, '1181-1257 <HAR>
 R/Wolfe, J.M.; Frank, R.; Muijo, K.; Spiro, R.C.; Reisfeld, R.A.; Rathjen, F.G.
 J. Biol. Chem. 263, 11943-11947, 1988
 A/Title: A human brain glycoprotein related to the mouse cell adhesion molecule L1.
 A/Reference number: A31072; MUID:8829876; PMID:3136168
 A/Accession: A31072
 A/Molecule type: protein
 A/Residues: 'O', '21-36 <WOL>
 R/Plutzer, M.; Bauer, D.; Drecher, B.
 submitted to the EMBL Data Library, March 1995
 A/Reference number: H01368
 A/Accession: G02506
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 11257 <PLA>
 A/Cross-references: EMBL:U52112; NID:g1302657; PIDN:AA051746.1; PID:g1302658
 C/Genetics:
 A/Gene: GDB:L1CAM
 A/Cross-references: GDB:120133; OMIM:303350; OMIM:308840
 A/Map position: Xq28-Xq28
 A/Intron: 26/1, 31/1; 66/2; 134/1; 175/1; 232/1; 269/2; 331/1; 375/1; 423/1; 460/2; 516
 /2
 C/Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology; 1
 C/Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; transmembran
 F/1-19/Domain: signal sequence #status predicted <SIG>
 F/20-1257/Product: neural cell adhesion molecule L1 #status predicted <MAT>

F/257-314/Domain: immunoglobulin homology <IMM1>
 F/532-593/Domain: immunoglobulin homology <IMM2>
 Query Match 5.6%; Score 131; DB 1; Length 1257;
 Best Local Similarity 19.2%; Pred. No. 1.2;
 Matches 92; Conservative 66; Mismatches 192; Indels 130; Gaps 21;
 QY 8 RHLLV-----LQALLPA--ATQGNKVLGKGDYELCTASQKSIQPHWKNQDKI 61
 DB 409 RHGLLANAYIVVQLPAILTADNQTVAVQGSTVYLCKFGAPVPSVQWLDGGTTV 468
 QY 62 LQNGSFLTKGSPKLNDRADSRSLWDQGNFPIIKNLKIEDSDTYICEVEDQKERVQL 121
 DB 469 LDQERF-----PYANGTIGTDLDQANDTGRFCLAAQNNVTIM 509
 QY 122 VFGLTANSDTHLIQ-----GGLTLTLES--PPGSSPSVQCSPPGKNIQ----- 164
 DB 510 A-NLKVKADATQTQGRSTTEKKSIVFTQNSFPSPSIQPSITWRGD--GRDLQELGDS 567
 QY 165 -----GKTLVSQLELOPSGTCTVLDQNKVKEKIDIVLAF-----QKASIVYKK 214
 DB 568 KYFIEDGR-LVHSLDYSQGNYSVASTELDVESRAQLLVGSPGVPRLVLSDLHL 626
 QY 215 EGRVFEFPPA-----FTVEKLTSGELMWQERASSRSKSWITFDKKNKESV 263
 DB 627 TQSQVAVSSPAEDNAPLEKIDIEFEDKEMAEKYSIGKPGNOSTITLKSPVYHT 686
 QY 264 KRYTODPKLQMGKPLHLTL--PQALPQ-----YAGSNTLTALBAKTGKLHQEVLVY 316
 DB 687 FRTALINKYGPSPVSEFVTVPAAPEKNPVDVQEGNET-----TNMYI 733
 QY 317 MRATQLQKLTCEWCP-----TSPKMLSLKENKAKVSKREPVVLAPEAGMWQ- 369
 DB 734 -----TWKPLRMMDMNAPOVYRVQ-----WRQSTRPWOE 765
 QY 370 CLTDSGVLTBSNIVLPTWSTPVHPRASALPAPPTGALPPOTASALPDPASALP 429
 DB 766 QIVSD--PFLVVSNTSTF---VPEIKVQAVNSGCK--DEPQVTTGSGEDYQALP 816
 RESULT 28
 T20992
 hypothetical protein F15G9.4a - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #ext change 18-Feb-2000
 C/Accession: T20992; T24733
 R/Suleton, J.
 submitted to the EMBL Data Library, December 1994
 A/Reference number: Z19355
 A/Accession: T20992
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-5175 <WLS>
 A/Cross-references: EMBL:Z47068; PIDN:CAA87335.1; GSPDB:GN00028; CESP:F15G9.4a
 A/Experimental source: clone F15G9
 R/Kerhew, J.
 submitted to the EMBL Data Library, December 1994
 A/Reference number: Z19929
 A/Accession: T24733
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-5175 <W12>
 A/Cross-references: EMBL:Z47070; PIDN:CAA87344.1; GSPDB:GN00028; CESP:F15G9.4a
 A/Experimental source: clone T09B9
 C/Genetics:
 A/Gene: CESP:F15G9.4a
 A/Map position: X
 A/Intron: 85/1; 120/1; 334/3; 370/1; 477/2; 606/3; 664/1; 935/3; 977/1; 1051/3; 1184/3;
 ; 2512/2; 2593/3; 2759/1; 2852/1; 2889/3; 2913/3; 2941/1; 2967/3; 2991/3; 3033/1
 ; 4225/1; 4361/1; 4408/1; 4456/1; 4498/1; 4647/3; 4838/1; 4879/1; 4941/1; 5011/1; 5077/
 Query Match 5.6%; Score 131; DB 2; Length 5175;
 Best Local Similarity 19.5%; Pred. No. 7.8;

Matches 85; Conservative 68; Mismatches 195; Indels 88; Gaps 18;

QY 33 KGDVETLTCTASQKSIQFHWKNSNQIKILNQGSLFTKGPCKLNDRADSRSLMDQGNF 92
 Db 3495 ENDITMDCGVTSRPLPSISM-----FRGDKPYLYD---RYSISPDGSH 3536

QY 93 PLIINKLKIEDSDTYICEVEDQ--KEVQLLVFGLT-ANSTHLLQGQSL-----TLTLE 144
 Db 3537 -ITINKAKLSDGKXICRASNEAGTSDIDLILKILVPRKIDKSNIIIGNPLAVARTIYLE 3595

QY 145 SPFGSSP-----SVQCRSPRGKNIQGGKTLVSQLELDQSGTCTVQLONQKV 193
 Db 3596 CPISGIPDPDVITWTKNGMDINMTDSRVLLAQNNEFFGLENVQVTDQGRYTCATNRGGA 3655

QY 194 EFKIDIVLAFQK---ASSIVYKKEGEOVERFPLAFTVEKLTGSGELMWQAERASSK 249
 Db 3656 SHDPSLDVLSPEFDHIGQPTIKRGGDTITLTCPKLAEDIADQVMDVSWTKDSRA--- 3712

QY 250 SWITFDL-KNKEVSVKRVTDPRKLQMGKKLPLHLTLPOALPOYAGSGNLTALAEKTKL 308
 Db 3713 --LDGDLTDNDVIS-----DDGRK---LTISQASLENNAGLYTCIALNRAGEASL 3756

QY 309 HOEVNVLVWRATQLOKN-----LTCEWVGPTSPKLMSLK-----LENKEAK 350
 Db 3757 EFKVEILSPVIDISRNDVQPOVAVNQPTIMCAVTHGHPFSIKM-LKNGKEVTDENIR 3815

QY 351 VSKREKPYVWLNP---AGMOCCLL-SDSGVLLLESNIKVLPTWSTPVPRASALPAPT 406
 Db 3816 IVEGQGVQLRLTSDSHAGKSCVAENDAGVKELE---MVLDFPTPVVSVKSDNPICAL 3872

QY 407 GSALPDPTASALPDP 422
 Db 3873 GETITLFCNAGSNPYR 3888

RESULT 29
 T43290
 hemicentin precursor - Caenorhabditis elegans
 C.Species: Caenorhabditis elegans
 C.Date: 11-Jan-2000 #sequence revision 11-Jan-2000 #text_change 18-Feb-2000
 C.Accession: T43290; T20993; T24734
 R.Vogel, B.E.; Hedgecock, E.W.
 submitted to the EMBL Data Library, June 1998
 A.Description: Hemicentin is required for hemidesmosome mediated cell adhesion and germ-
 A.Reference number: 222396
 A.Accession: T43290
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: mRNA
 A.Residues: 1-5198 <VOG>
 A.Cross-references: EMBL:AF074901; PIDN:AAC26792.1
 R.Sulton, J.
 submitted to the EMBL Data Library, December 1994
 A.Accession: T20993
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-5198 <WID>
 A.Cross-references: EMBL:Z47068; PIDN:CAA87336.1; GSPDB:GN00028; CESP:FL5G9.4b
 A.Experimental source: clone F15G9
 R.Kershaw, J.
 submitted to the EMBL Data Library, December 1994
 A.Reference number: Z19929
 A.Accession: T24734
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-5198 <WID>
 A.Cross-references: EMBL:Z47070; PIDN:CAA87345.1; GSPDB:GN00028; CESP:FL5G9.4b
 A.Experimental source: clone T09B9
 C.Genetics:
 A.Gene: him-4; F15G9.4b
 A.Map position: X
 A.Introns: 85/1; 120/1; 334/3; 370/1; 477/2; 606/3; 664/1; 935/3; 977/1; 1051/3; 1184/3;
 ; 2512/2; 2593/3; 2699/3; 2759/1; 2852/1; 2889/3; 2913/3; 2941/1; 2967/3; 2991/3; 3033/1

1; 4225/1; 4361/1; 4408/1; 4456/1; 4498/1; 4647/3; 4838/1; 4902/1; 4964/1; 5034/1; 5100/

Query Match 5.6%; Score 131; DB 2; Length 5198;
 Best Local Similarity 19.5%; Pred. No. 7.9;
 Matches 85; Conservative 68; Mismatches 195; Indels 88; Gaps 18;

QY 33 KGDVETLTCTASQKSIQFHWKNSNQIKILNQGSLFTKGPCKLNDRADSRSLMDQGNF 92
 Db 3495 ENDITMDCGVTSRPLPSISM-----FRGDKPYLYD---RYSISPDGSH 3536

QY 93 PLIINKLKIEDSDTYICEVEDQ--KEVQLLVFGLT-ANSTHLLQGQSL-----TLTLE 144
 Db 3537 -ITINKAKLSDGKXICRASNEAGTSDIDLILKILVPRKIDKSNIIIGNPLAVARTIYLE 3595

QY 145 SPFGSSP-----SVQCRSPRGKNIQGGKTLVSQLELDQSGTCTVQLONQKV 193
 Db 3596 CPISGIPDPDVITWTKNGMDINMTDSRVLLAQNNEFFGLENVQVTDQGRYTCATNRGGA 3655

QY 194 EFKIDIVLAFQK---ASSIVYKKEGEOVERFPLAFTVEKLTGSGELMWQAERASSK 249
 Db 3656 SHDPSLDVLSPEFDHIGQPTIKRGGDTITLTCPKLAEDIADQVMDVSWTKDSRA--- 3712

QY 250 SWITFDL-KNKEVSVKRVTDPRKLQMGKKLPLHLTLPOALPOYAGSGNLTALAEKTKL 308
 Db 3713 --LDGDLTDNDVIS-----DDGRK---LTISQASLENNAGLYTCIALNRAGEASL 3756

QY 309 HOEVNVLVWRATQLOKN-----LTCEWVGPTSPKLMSLK-----LENKEAK 350
 Db 3757 EFKVEILSPVIDISRNDVQPOVAVNQPTIMCAVTHGHPFSIKM-LKNGKEVTDENIR 3815

QY 351 VSKREKPYVWLNP---AGMOCCLL-SDSGVLLLESNIKVLPTWSTPVPRASALPAPT 406
 Db 3816 IVEGQGVQLRLTSDSHAGKSCVAENDAGVKELE---MVLDFPTPVVSVKSDNPICAL 3872

QY 407 GSALPDPTASALPDP 422
 Db 3873 GETITLFCNAGSNPYR 3888

RESULT 30
 S04663
 T-cell receptor gamma chain - human (fragment)
 C.Species: Homo sapiens (man)
 C.Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text_change 21-Jan-2000
 C.Accession: S04663
 R.Okada, A.; Bank, I.; Rogozinski, L.; Takihara, Y.; Mak, T.W.; Cheese, L.; Alt, F.W.
 J. Exp. Med. 168, 1481-1486, 1988
 A.Title: Structure of the gamma/delta T cell receptor of a human thymocyte clone.
 A.Reference number: S04663; MUID:89010543; PMID:2844954
 A.Accession: S04663
 A.Molecule type: DNA
 A.Residues: 1-304 <OKA>
 A.Cross-references: EMBL:X15018
 A.Note: The authors translated the codon AAC for residue 72 and AAT for residues 102 and
 C.Genetics:
 A.Introns: 15/1; 133/3
 C.Superfamily: immunoglobulin V region; immunoglobulin homology
 C.Keywords: T-cell receptor
 F.158-223/Domain: immunoglobulin homology <IMM>

Query Match 5.6%; Score 130.5; DB 2; Length 304;
 Best Local Similarity 21.3%; Pred. No. 0.19;
 Matches 74; Conservative 59; Mismatches 128; Indels 87; Gaps 15;

QY 8 RHLLVLQALLPATQGNKVVLGKGDY-----VELTCTASQKSIQFHW----- 53
 Db 2 RMAILLVLAFLSPASQSSNLEGGTKSVTRPRSSAEITCDLTVINAFYIMHYLHQEGKA 61

QY 54 -----KNSNQIKILG-NQGSFLTGPCKLNDRADSRSLMDQGNFPLIKYKTED 103
 Db 62 PQRLLYDVNSKQVLESGLSPKYYHTP-----KRMSW-----ILLRLIEND 107

QY 104 SDTYICEVEDQKEVQLLVFGLTANSDTHLLQGQSLTLT--LSPFGSSPSVQCRSPRG 160

```

Db      108 SGVYCATMDRQK---KLFG-----SGTTLVTDQKDADADVSPKTFILP----- 150
QY      161 KNIGQGKTLVSQLELDSDGTWCTYQLQNKQKVEFKIDIVLAFQ--KASSIYKKEGEG 218
Db      151 -----SIATKTKQKAGTYLCL-----SKFPDIIKIKMOEKSKSTILIGSEGN 195
QY      219 -----VEFSPLAFVETKLTGSGELMWQAEASSSKSWITFDLKNKEVSRYKVTQDPK 271
Db      196 MKTNDTYMKFSW-LIYPEESLDENHRC---IVHNNKNGIDDEIIFPPIKTDVTTVDPK 251
QY      272 LQWKKLPLHLTPQALPOYAGSGNLTALAEAKTGKLDHENVLVMRA 319
Db      252 YNYSKQANDVITM-DPKDMSKQANDTLTLQLNTSAIYVYLLLLKS 298

RESULT 31
myosin-light-chain kinase (EC 2.7.1.117), 210K, nonmuscle - chicken
N:connectin; myosin-light-chain kinase, 108K, smooch muscle; telokin
C:Species: Gallus gallus (chicken)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S68235; A37099; B4389; S28227; S78216; A35093; A25810; S11652
R:Walterson, D.M.; Collinge, M.; Lukas, T.J.; Van Eldik, L.J.; Bitukov, K.G.; Stepanova,
FEBS Lett. 373, 217-220, 1995
A>Title: Multiple gene products are produced from a novel protein kinase transcription
A:Reference number: S68235; MUID:96033976; PMID:7589469
A:Accession: S68235
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1906 <EMBL>
A:Cross-references: EMBL:X52876; NID:9992992; PIDN:CAA37056.1; PID:9992993
R:Shoemaker, M.O.; Lau, W.; Shattuck, R.L.; Kwiatkowski, A.P.; Matrisian, P.E.; Guerra-S
J. Cell Biol. 111, 1107-1125, 1990
A>Title: Use of DNA sequence and mutant analyses and antisense oligodeoxynucleotides to
activity.
A:Reference number: A37099; MUID:90361738; PMID:2202734
A:Accession: A37099
A:Molecule type: mRNA
A:Residues: 649-1906 <SHO>
A:Cross-references: EMBL:X52876
R:Collinge, M.; Matrisian, P.E.; Zimmer, W.E.; Shattuck, R.L.; Lukas, T.J.; Van Eldik, I
Mol. Cell. Biol. 12, 2359-2371, 1992
A>Title: Structure and expression of a calcium-binding protein gene contained within a c
A:Reference number: A44389; MUID:92236611; PMID:1373815
A:Accession: B44389
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1695-1906 <COL>
A:Cross-references: GB:M88284; NID:g212237; PIDN:AAB5767.1; PID:g212238
A:Accession: A44389
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1750-1906 <CO2>
A:Cross-references: GB:M88283; NID:g211371; PIDN:AAA48647.1; PID:g211372
R:Yoshikai, S.I.; Ikebe, M
Arch. Biochem. Biophys. 299, 242-247, 1992
A>Title: Molecular cloning of the chicken gizzard telokin gene and cDNA.
A:Reference number: S28227; MUID:93073972; PMID:1444462
A:Accession: S28227
A:Molecule type: mRNA
A:Residues: 1750-1906 <YOS>
A:Cross-references: EMBL:M6655; NID:g212744; PIDN:AAA49083.1; PID:g212745
A:Accession: S78216
A:Molecule type: DNA
A:Residues: 1750-1906 <YOM>
A:Cross-references: EMBL:M69687
R:Olson, N.J.; Pearson, R.B.; Needleman, D.S.; Hurwitz, M.Y.; Kemp, B.E.; Means, A.R.
Proc. Natl. Acad. Sci. U.S.A. 87, 2284-2288, 1990
A>Title: Regulatory and structural motifs of chicken gizzard myosin light chain kinase.
A:Reference number: A35093; MUID:90192792; PMID:2315320
A:Accession: A35093
A>Status: preliminary

```

```

A:Molecule type: mRNA
A:Residues: 935-1438, 'Q', 1440-1906 <OLS>
A:Cross-references: GB:M31048; NID:g212660; PIDN:AAA9069.1; PID:g212661
R:Guerrero Jr., V.; Russo, M.A.; Olson, N.J.; Putkey, J.A.; Means, A.R.
Biochemistry 25, 8372-8381, 1986
A>Title: Domain organization of chicken gizzard myosin light chain kinase deduced from
A:Reference number: A25810; MUID:87157587; PMID:3030394
A:Accession: A25810
A:Molecule type: mRNA
A:Residues: 1258-1438, 'Q', 1440-1906 <GUE>
C:Genetics:
A:introns: 1735/3; 1779/1; 1819/1
C:Superfamily: myosin-light-chain kinase, nonmuscle; fibronectin type III repeat homolo
C:Keywords: alternative initiators; ATP; calmodulin binding; phosphoprotein; phosphotran
F:542-599/Domain: immunoglobulin homology <IMM1>
F:935-1906/Domain: myosin-light-chain kinase, 108K, smooch muscle (from 5.5kb transcript
F:1098-1158/Domain: immunoglobulin homology <IMM2>
F:1451-1708/Domain: protein kinase homology <KIN>
F:1459-1467/Region: protein kinase ATP-binding motif
F:1750-1906/Product: telokin (kinase-related protein KRP) (from 2.7 kb transcrip
F:1808-1869/Domain: immunoglobulin homology <IMM3>

```

Query Match 5.6%; Score 130; DB 1; Length 1906;

Best Local Similarity 20.8%; Pred. No. 2.4;

Matches 82; Conservative 58; Mismatches 142; Indels 112; Gaps 15;

89 QGNFPIITNKLIEDSDTYICEVD---QKEVQLVFGLTANSPTHLQGSITLTL 144

81 RGIFSLVIRGVQEGSGKTYCEANDGVRQTVETLVEG---NS---LKKYSLPSSAK 133

145 SPFG---SSPSVOCR-----SPRGNIGGKT----- 168

134 TREGRLSPVPEHRPSIMGESPPKATKPNRVVBEQGTGRFSKITGRPPQVQVTWKCD 193

169 -----LSVSOLELDSDGTWCTYQLQNKQKVEFKIDIV----- 201

194 IHLQNERFMPEKTIQVLEIQNVQLDAGITCTVNVSAKASVSLVYQGPDKTD 253

202 -----LA-----FOKASSIYKKEGVEFSPLAFYEKLTSGSEL 240

254 HAOPLCMPKPTTLTKALINSDFKOATNSNGIAKEIKSTEL-VWETDRLSAKKEPFY 312

241 QAEBASSSKSWITFDLKNKEVSKRVTQDPKQLQWKKLPLHLTPQALPOYAGSGNLT 300

313 TSEBAKQKQGNQANAVPLQESRGTGKPPVLQKTSSTITLQAVVAQDEPPAEPQTFI 372

301 LEAKTGKLDHENVLVMRATQLQKNLTCEVWGPTSPKMLSLKEN-----KEAKYSKE 355

373 RQAEERK--RTVQPLMTTITQENPSLT-----QGVSPR---SRETFNRAGVRKSVBEKKE 423

356 KPVWVNLNPEAGMWQCLLSDSQ-VLLESNIKVL 388

424 -PLGIPPGFESRPGQLEASGGEIRKFSKVSQGP 456

RESULT 32

T42633 connectin/titin - chicken (fragment)

C:Species: Gallus gallus (chicken)

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000

C:Accession: T42633

R:Ujima, H.; Ohtsuka, H.; Kawamura, Y.; Kume, H.; Murayama, T.; Abe, H.; Kimura, S.; Ma

Biochem. Biophys. Res. Commun. 223, 160-164, 1996

A>Title: A 11.5-kb 5'-terminal cDNA sequence of chicken breast muscle connectin/titin re

A:Reference number: Z22221; MUID:96254045; PMID:8660363

A:Accession: T42633

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-4162 <YAJ>

A:Cross-references: EMBL:D83390; NID:g1513029; PIDN:BA11908.1; PID:g1513030

A:Experimental source: breast muscle

C:Keywords: skeletal muscle

```

Query Match          5.6%; Score 130; DB 2; Length 4162;
Best Local Similarity 22.4%; Pred. No. 6.8;
Matches 103; Conservative 55; Mismatches 159; Indels 142; Gaps 24;

QY 11 LVLQALLPAATQGNKVVLGKKDVELTCTASQKKSIOFW-KNSNOI-----KILG 63
DB 3551 LTVLEPAVIVEKPGPVKVT---AGDSCTLECTVDGTPELTARWFDGMLSTDHKKYKI-- 3605
QY 64 NGGSFLTK--GPSKLNDRADSRSLMDGNFPLITIKIEDSDPYICVEVDQKEVQL 121
DB 3606 ---SFPNKSGLKILN-----AGLEDSEYTFEYKNSGKSSCTASLQVSDR----- 3649
QY 122 VEGLTANSDTHLQGSLLTLE---SPGSSPSVQCR--SP-----RGKNIQGGK 167
DB 3650 -----IMPSFTRKLEKETYGLQSSAVLECKVYGSPPILVSWFHDGQETSGD 3697
QY 168 -----TLSSQLELQDSGTMCTVLOKKNVEKIDIVLAPQKASIIYVKKK- 215
DB 3698 KYQATLTDTNCTGLKNGQESDMGTSTATVAVSGSCS--AFLSVAREPPSPVKKEP 3754
QY 216 -----GEOVEFSFPLAFTVEKLTGSGELW---QAERASSKSWITFDLKNKEVSVKRV 267
DB 3755 FNVLSGENTFTT-----SIVKGSPLLEVNMFRGSIETLAFGHKCNITL----- 3796
QY 268 QD--PKLQNGKKLPLHLTLPOALPOYAGSGNLTALAEATGKLHQEVLN-----VYMR 318
DB 3797 QDSVALELFDVQPLQ-----SGDYTCQVSNBAGKISCTTHLFLVKEPAKFWK 3844
QY 319 ANQQLQ---KNLTCEFWGPTSEKMLSLKLEKKEAKVSKREKPVWLVNEA----- 365
DB 3845 VNDLSVEKKNLILECTYTGTPTISATWK--KNGVILHSSKCSITTTETSAIIETPNS 3901
QY 366 -----GMQOC-LTSDSGVLTLSNLIKVL--PTMSTFVHP 396
DB 3902 KLEDGQVYSCHLENDSGQDNCGATITILEPFFVTPLEP 3940

RESULT 33
T32735
telomerase-associated protein component 1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C:Accession: T32735
R:Nakayama, J.; Saito, M.; Nakamura, H.; Matsura, A.; Ishikawa, F.
Cell 88, 875-884, 1997
A:Title: TLPI: a gene encoding a protein component of mammalian telomerase is a novel me
A:Reference number: Z21217; MUID:97236507; PMID:9118230
A:Accession: T32735
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2629 <NAK>
A:Cross-references: EMBL:U89282; NID:91932816; PID:91932817; PIDN:ABS1690.1
A:Experimental source: strain Fischer 344
A:Genetics:
A:Gene: TLPI
C:Function:
A:Description: modification of telomerase-associated protein component 1 may regulate te

```

```

QY 139 LTL-----TLESPGSSPSVQCRSPRGKN-----IQGKTLVSLQLELQDS 179
DB 2395 LILNKEDVELLQMKRGSTPSSICRVAHSSILCTSKDYGLFYIQGNGSSLSLEQES 2454
QY 180 GTWCTCTVLOKQKVEFK-----IDIVLAPQKASIIYVKKKEQVEFSPLAFTV 229
DB 2455 GKF-----EKTIDFNLNLNPNNGSPVSIQAEPEGSSILL-----CATSDGMLMNL 2500
QY 230 EKLQSGE-----LWQAERASSSK-----SWITFDLKNKEVSVKRVQDP 270
DB 2501 SECTPEGEWVVDNLTWQKSRNPKSRTTPTDSSPGLFCMDSNV-----BPTHILARCK 2553
QY 271 KLGQKKLPLHLTLPOALPOYAGSGNLTALAEATGKLHQEVLNVMRATQLOKNTLC-E 329
DB 2554 KIHGSVTALHV-----LP-----GLVLTASEBDVYLMERPMQGLGLPRCEGVSCLE 2603
QY 330 VWGFTSPKLMUSL 342
DB 2604 PMWEPSPQLQAV 2616

RESULT 34
S18252
heparan sulfate proteoglycan - mouse
N:Alternate names: perlecan
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C:Accession: S18252; A31917; B31917; S66460
R:Noonan, D.M.; Fuller, A.; Valente, P.; Cai, S.; Horigan, E.; Sasaki, M.; Yamada, Y.; Ha
J. Biol. Chem. 266, 22939-22947, 1991
A:Title: The complete sequence of perlecan, a basement membrane heparan sulfate proteogly
adhesion molecule.
A:Reference number: S18252; MUID:92078153; PMID:1744087
A:Accession: S18252
A:Molecule type: mRNA
A:Residues: 1-3707 <NOC>
A:Cross-references: EMBL:M77174; NID:9200295; PIDN:AAA39912.1; PID:9200296
R:Noonan, D.M.; Horigan, E.A.; Ledbetter, S.R.; Vogeli, G.; Sasaki, M.; Yamada, Y.; Ha
J. Biol. Chem. 263, 16379-16387, 1988
A:Title: Identification of cDNA clones encoding different domains of the basement membra
A:Reference number: A92680; MUID:89034110; PMID:2972708
A:Accession: A31917
A:Molecule type: mRNA
A:Residues: 940-1601 <NO2>
A:Cross-references: GB:J04054; NID:9200252; PIDN:AAA39899.1; PID:9200253
A:Accession: B31917
A:Molecule type: mRNA
A:Residues: 1870-2600 <NO3>
A:Cross-references: GB:J04055; NID:9200300; PIDN:AAA39912.1; PID:9200301
R:Schulze, B.; Mann, K.; Battistutta, R.; Wiedemann, H.; Timpl, R.
Eur. J. Biochem. 231, 551-556, 1995
A:Title: Structural properties of recombinant domain III-3 of perlecan containing a glob
A:Reference number: S66460; MUID:95377282; PMID:7649134
A:Accession: S66460
A:Molecule type: Protein
A:Residues: 1272-1274, 'X', 1276, 'X', 1278-1279 <SCH>
C:Superfamily: LDL receptor ligand-binding repeat homology; BGF homology; laminin G repe
C:Keywords: glycoprotein
F:1199-234/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:1285-319/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:325-359/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:368-403/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:764-811/Domain: laminin-type EGF-like homology <LEG>
F:1159-1206/Domain: laminin-type EGF-like homology <LEG7>
F:11563-1610/Domain: laminin-type EGF-like homology <LEG8>
F:16163-1668/Domain: laminin-type EGF-like homology <LEG9>
F:3163-3198/Domain: BGF homology <EGF>
F:3270-3423/Domain: laminin G repeat homology <LG2>
F:3464-3492/Domain: BGF homology <EGF7>
F:1256,1891,2336,2394,2427/binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match          5.5%; Score 127; DB 2; Length 3707;

```

Best Local Similarity 19.3%, Pred. No. 9.2;
Matches 100; Conservative 66; Mismatches 185; Indels 168; Gaps 22;

```
Oy 32 KGGDTVEITCTASOKKSIQFHKNSNOIKILGNQSEFLTKGPKSLNDRADSRSLMDQGN 91
Db 2447 KMGKDITLLECISSEBPRSSPRMTR-----LGI-----PYLLEPRMGKLN-----SH 2488
Oy 92 PFLIIKOLKIEDSDTYICEVED---OKEEVQLLVGLTANSDTHLLIQGSLTLTLTSP 147
Db 2489 AMLKIASKPSDAGTAVCOAQNALGTAKQVELIVDTGVAPGTPQVVESEELTEA-- 2546
Oy 148 GSSPSVQCRS-----PRGNIQGKTLISQLELDDSGTWTCTVQLNQ 150
Db 2547 GHTYTLHCASAGNDPPTIHWSKLAPLPWQRIE-GNTLVIPRAQDSDQYICNATNSA 2605
Oy 191 KKEVFKIDIVLAFQKASSI---VYKKEGEQVESEF-----PLAF----- 227
Db 2606 GHTBATVVLNHSPPRYATILEHTSAQPNLVQCLAHGTPPLTYQMSLVGVLPKAV 2665
Oy 228 -----TYEKLIGSGELMWQAERASSKSWITFDLKNKEVSKRV 266
Db 2666 VRNQLRLRLETPVPEDSGKRVCGVSNRVSABAFQVLVQSSSNLPTDTSIPGSGTPVQV 2725
Oy 267 TDPRKLO---MGKLLPLHLTPQA---LPRYASSGNLTALAEKTKGLKAEVNLVYMA 319
Db 2726 T-POLETRNIGASVEFHCAVPNERGTHLRWLKEGGOLPRGHSVQDG-----VLRI 2774
Oy 320 TOLQKN---LTCEWGP-----TSPKMLSLKLEKAKV----- 351
Db 2775 QNLQNCQGTIVVCOAHGFWGOAQATQALIVQALPSVLINRTSVHSVYVGHSEFECAL 2834
Oy 352 -SKREKPVV-----LNP-----EAGMNOCLSDS-----GVLLSN 383
Db 2895 GDPKQVYTWSKVGGHLRGIVQSGTIRIRIAHVELADAGQVRCATNAAGTTQSHVLL--L 2892
Oy 384 IKVLPWSTPVNPRASALPAPPTGSLPDPGTASALPDP 422
Db 2893 VQALPQISTPREIRV-----PAGSAVFPFCMASGYPRP 2925
```

RESULT 35
JUMSWG
neural cell adhesion molecule 1 precursor, GPI-anchored splice form - mouse
N:Alternate names: NCAM-120
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text_change 31-Dec-2000
C:Accession: A29673; S00382; A44290
R:Bartheleis, D.; Sancio, M.D.; Wille, W.; Ruppert, C.; Chaix, J.C.; Hirsch, M.R.; Fontec
EMBO J. 6, 907-914, 1987
A:Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000
A:Reference number: A29673; MUID:87246524; PMID:3595563
A:Accession: A29673
A:Molecule type: mRNA
A:Residues: 1-725 <BA2>
A:Cross-references: EMBL:Y00051; NID:G53342; PIDN:CA68263.1; PID:G53343
R:Barbae, J.A.; Chaix, J.C.; Steinmetz, M.; Goriidis, C.
EMBO J. 7, 625-632, 1988
A:Title: Differential splicing and alternative polyadenylation generates distinct NCAM
A:Reference number: S00382; MUID:88283628; PMID:3396534
A:Accession: S00382
A:Molecule type: DNA
A:Residues: 642-656, 'D', 658-725 <BA2>
A:Cross-references: EMBL:X07195
R:Rougon, G.; Marhak, D.R.
J. Biol. Chem. 261, 3396-3401, 1986
A:Title: Structural and immunological characterization of the amino-terminal domain of m
A:Reference number: A44290; MUID:86140120; PMID:3512556
A:Accession: A44290
A:Molecule type: protein
A:Residues: 20-36 <ROU>
C:Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol
C:Comment: Several forms of NCAM are produced by alternative splicing. See also PIR:JUMS
C:GeneID:3

A:Gene: NCM
A:Map position: 9
A:Initrons: 701/1
C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immunoglobulin domain
C:Keywords: alternative splicing; cell adhesion; duplication; heparin binding; membrane
F:1-19/Domain: signal sequence #status predicted <Sig>
F:13-98/Domain: immunoglobulin homology <IMM1>
F:132-191/Domain: immunoglobulin homology <IMM2>
F:152-156/Region: heparin binding #status predicted
F:161-165/Region: heparin binding #status predicted
F:228-290/Domain: immunoglobulin homology <IMM3>
F:263-272/Region: NCM binding #status predicted
F:323-388/Domain: immunoglobulin homology <IMM4>
F:420-482/Domain: immunoglobulin homology <IMM5>
F:519-596/Domain: fibronectin type III repeat homology <FN3a>
F:625-685/Domain: fibronectin type III repeat homology <FN3b>
F:741-961/Domain: fibronectin type III repeat homology <FN3c>
F:1222,316,348,424,450,479/binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.4%; Score 125.5; DB 1; Length 725;
Best Local Similarity 20.9%; Pred. No. 1.3;
Matches 95; Conservative 64; Mismatches 187; Indels 109; Gaps 21;

QY 3 RG-VPRHLLVLTQALLPAAATGKAVYLGGKEDYELTSTAQKSIQPHW-KNSNQIK 60
DB 196 RGEINFDIIVVNPPTVQARQSIYVATNLIQSVTLVVDADGFPPTSMWKDGEPIE 255
QY 61 ILNGQSFLLTKGPSKLNDRAD--SRRLMDQGNFPLIKYLIKLEDDYICEVE---DQ 114
DB 256 -----NEEDERSRSSVDSSE--VTIRVVDKNDDEAYCIENKAGEQ 257
QY 115 KEVQVLVFG---LTANSDTHLLQGSLLTLLSPGSSPSVQCR-----SPRGKNIQ 164
DB 298 DASIHLKVFAPKPTIYVENQTALEBERQVLTTCBASODPIPSITMRSTRNISSEBDQ 357
QY 165 GG-----KTLVSQLELDPSGTWTCTVLDQKKVKEKIDIVLAFAKASI----- 210
DB 358 GHVYVRSARVSSILTKLSIGYRDAGEYTCASNTIGDQSQID---LEFYAKLQGPVA 414
QY 211 VYKKEGQVE-----FSFPLAFTVEKLTGSGELMMQERASSSKSWITPDLKA-KEVSYK 264
DB 415 VYTBEGQVAVITCEGVAFYPSA-TIS-----WFRDQGLLPBSNYSNITKIYTPPSAYL 465
QY 265 RVTQDPPLQMGK-----KLPLHLTLPLQ-----LPQAGSGNLTLLAEK 304
DB 466 EYTPDSSNDPQGNVNCATVNRNIGQESLEFLLVQADPTSPSIDRVEPPSYAQVQDPEPA 525
QY 305 TGLKHQEVNLVMPRATOLQNLTCBEVWGPTSPKMLSLKLENKYAKYSKREKVVWLNPE 364
DB 526 TG-----GVPILIKYKAEMKSLGESVHFT---WYDAKANNMEGIYV-----IMGLKPE 570
QY 365 AGMWQCLLSDSGVLT-----LESNIKVLPTWSP 393
DB 571 TTYSDRLLAALNGKGLGEIMQPSSESKTPQVPELSAP 605

RESULT 36
A:54100
A:Residues: 1-1447 <HE>
A:Residues: 1-1447 <HE>
A:Cross-References: EMBL:X6132; NID:g453209; PIDN:CAAS3735.1; PID:g453210
A:Pearson, E.R.; Cho, K.R.; Nigro, J.M.; Kern, S.E.; Simons, J.W.; Ruppert, J.M.; Hamilton
science 247, 49-56, 1990

A:Title: Identification of a chromosome 18q gene that is altered in colorectal cancers.

A:Reference number: A40098; MUID:90100559; PMID:2294591

A:Accession: A40098

A:Molecule type: mRNA

A:Residues: 1-750 <FEA>

A:Cross-references: GB:M32292; NID:G181492; PIDN:AAA3751.1; PID:G181493

C:Genetic8:

A:Gene: GDB:DC

A:Cross-references: GDB:119838; OMIM:120470

A:Map position: 18q21.1-18q21.1

C:Keywords: transmembrane protein; tumor suppressor

F:1-25/Domain: signal sequence #status predicted <SIG>

F:26-1447/Product: tumor suppressor protein DCC #status predicted <MAT>

Query Match 5.4%; Score 125.5; DB 2; Length 1447;

Best Local Similarity 19.3%; Pred. No. 3.3; Indels 147; Gaps 15;

Matches 77; Conservative 46; Mismatches 129; Indels 147; Gaps 15;

Qy 34 GDTVELTCTASQKSIQFHWKNSNQIKILGNQSFITKPSKLNDRADSRSLMDQGNFP 93

Db 154 GDTVLKCEVIGEPMTIHMQNQ-----DLTRIP-----GDSRVVLPSSG--A 196

Qy 94 LIINKLKIEDSDTYICEVED-----QKEVQLI-----VGLTANSDTHLQGS 138

Db 197 LOISRLQPGDIGYRCASRNPAASRTGNEAEKRLISDPGLHROLYFLORPNNVAIEGKD 256

Qy 139 LTL--TLESPPG-----SPSVQCRSPRGKNIQKGLTSVSOQLDSDGTMTCTVQN 189

Db 257 AVLECCVCGYPPPTFTLGRGEVITQKSK--YSLGGSNLILSNVTTDDSCMYTCVTVYK 315

Qy 190 QKVEFKIDIVLA---FQKASSIYKKEGQVEFSPFLAFTVEKLTGSGELMWOAERAS 246

Db 316 NENISASAEITLVLPWFLNHPNLSNLYAYESMDIE-----ECTVSG-----KPV 359

Qy 247 SSKSKITTDLKNKEYSVKRVTDPKLQNGKKLPLHLITLPQALPYAGSGLTALAEKATG 306

Db 360 PTVNM---MKNGDV-----VIPSDFYFQIVGSGNLR----- 387

Qy 307 KLHGEVNLVWRATQLOKNLTCCEWGPSPKMLSLKLENKAQVSKREKVVWVLNPEAG 366

Db 388 -----LGVVKSDEG 396

Qy 367 MMQCLSDSGOVLLESNIKVLPTWSTPVHPASALPAP 405

Db 397 FYQCVAEENAGNAQISQAQIIVP--KPAIPSSSVLPSPAP 432

RESULT 37

108851

Down syndrome cell adhesion protein 1 - human (fragment)

N:Alternate names: Down syndrome cell adhesion molecule

C:Species: Homo sapiens (man)

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999

C:Accession: T08851

R:Yamakawa, K.; Hsu, Y.K.; Haendel, M.A.; Hubert, R.; Chen, X.N.; Lyons, G.E.; Korenberg

submitted to the EMBL Data Library, September 1997

A:Description: DSCAM: A novel member of the immunoglobulin superfamily maps in a down sy

A:Reference number: Z16495

A:Accession: T08851

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1896 <YAM>

A:Cross-references: EMBL:AF023449; NID:G3169765; PID:G3169766

A:Experimental source: brain; developmental stage: 14 weeks; fetal

C:Genetic8:

A:Gene: DSCAM

A:Map position: 21q22

A>Note: derived from alternately-spliced mRNA

C:Function:

A:Description: involved in nervous system development

C:Keywords: alternative splicing

Query Match 5.4%; Score 125; DB 2; Length 1896;

Best Local Similarity 18.6%; Pred. No. 5.1;

Matches 82; Conservative 55; Mismatches 144; Indels 160; Gaps 17;

Qy 20 PATQGNKVVLGKKGDTVELTCTASQKSIQFHW-KNSNQIKILGNQSFITKPSKLNDR 78

Db 490 PASIRPKNKTITAJAGRDYTHCRVIGYPPYSIKRYKNSNLPNNHROVAFENNGTLKSD 549

Qy 79 RADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQEEVQLVGLTANSDTHLQGS 138

Db 550 VQ-----KEVDEGEYTCNVLPQ-----LSTQS 574

Qy 139 LTLTLESPP-----GSSPSVQCRSPRG-----KNT 163

Db 575 VHTVVKVPFIQPEPFPSIIGQRFVPCVVGSDLPITITWQKGRPIGSLGVITDNI 634

Qy 164 QGKGLTSVSOQLDSDGTMTCTVQLQNGKVEFKIDIVLA-----FQKASS 209

Db 635 DFTSSLRISMLSMHNGNYTCIARNEAAVEHQSOLIVRPKRVVQPRDQGIYGCAVI 694

Qy 210 IYVKEGEQVEFSPFLAFTVEKLT-GSGELMWOAERASSKSWITFDLKNKEYSVK----- 264

Db 695 LNCSAEG-----YVPFTYWKFSKAGVPOFO-PIALNGRIQV--LSNGSLIKHIVE 744

Qy 265 -----RYTQDPKLOMGKKLPLHLITLPQALPYAGSGLTALAEKATGKLHGEVNLV 316

Db 745 EDGQYVLCKVSNVDGADVSKMYLTIVKIPAMITSP--NTTLATQGO----- 789

Qy 317 MRATQLOKNLTCCEWGPSTFKMLSLKLENKAQVSKREKPV--W-----VLPENAGMW 368

Db 790 -----KKEMSCYAHG-----EKPIIVWEKEDRIITPEMAY 821

Qy 369 QCLSDSGOVLLESNIKVLPT 389

Db 822 LVSTKEVGEVIT-STQLIILPT 841

RESULT 38

S20901

ctln - rabbit (fragment)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Jun-1999

C:Accession: S20901; I46520

R:Label, S.; Barlow, D.P.; Gautel, M.; Lakey, A.; Trinick, J.

EMBO J. 11, 1711-1716, 1992

A:Title: A regular pattern of two types of 100-residue motif in the sequence of titin.

A:Reference number: S20897; MUID:92258380; PMID:1582406

A:Accession: S20901

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-6805 <LAB>

A:Cross-references: EMBL:X64696

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1992

R:Label, S.; Barlow, D.P.; Gautel, M.; Gibson, T.; Holt, J.; Heien, C.L.; Francke, U.; I

Nature 345, 273-276, 1990

A:Title: A regular pattern of two types of 100-residue motif in the sequence of titin.

A:Reference number: I46520; MUID:90238553; PMID:2129545

A:Accession: I46520

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 4235-5250 <IA2>

A:Cross-references: EMBL:X17329; NID:G1756; PIDN:CAA35207.1; PID:G930251

C:Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; prot

C:Keywords: muscle

Query Match 5.4%; Score 125; DB 2; Length 6805;

Best Local Similarity 18.6%; Pred. No. 28;

Matches 111; Conservative 81; Mismatches 180; Indels 224; Gaps 26;

Qy 11 LTVLQALPAA-TGG--NNVVLGKKGDTVELTCTASQKSIQFHWKNSNQIKILGNQGS 67

Db 2916 VIVKEQWLPEDLDLRGIVQGLVATAGADNITKEIIPVLRKPKPTVYWKSGQVQLKQTORVN 2975

Qy 68 FLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVED-----QKEEV 118

```

Db 2976 VENTATSTILINSECVRS--DSCGPRFLAKNI VGEVGVITTIQVNDIGPPTGPIKPEV 3033
Qy 119 QLVVFGLTANSPDTHLLQGOSLTLTLESPGSSPSVQSSPRKGIQGGKTS--VSQLEL 176
Db 3034 -----SSD-----FVTFPSWEPPEN-----DGVVPSISNVYIEKRO 3062
Qy 177 QDSGTV---TCTVLNOKK-----VEFKIDI-----VLAFOKXSSI----- 210
Db 3063 TDSITWELATVITVITTYKATRLTGTVEYQFRVKAQNRVGVPGITTSASIVANYPEKVG 3122
Qy 211 -----VYKKEGQVEFSFPLA-----FTVEKLTGSGELMWQAEARA----- 245
Db 3123 PGRTQVTAVTKDSMTISGHEPLSDGSGPILOHYERKERKRNILMQTVSKALVPENIFKS 3182
Qy 246 SSSKSWITFDL-----KXKEVSKRVTDQPKLQMGKPLHL-----TLPOALP 289
Db 3183 SGLTDGIAYEPRVIAENNAKSKPSKPSBPVLDLPIDPPGKPIPLNITRHTVTLKMAKP 3242
Qy 280 QVAGSGNLTLMLAK-----TGKLNQ-----EVNLVVRANQ-----LQKXLTCEV 330
Db 3243 ETGGGFKITSYIVKRDLPNGRWLKAQNSNIENFTVSGLTEDAAVEPRVIAKNAAGAT 3302
Qy 331 WGPV-----SPKMLSLKLEBK-----EAKYSKEEKP----- 357
Db 3303 SPSPSPDAITCRDVEAPRILVDVRFKDTVLKAGFAFKLEADVSGRPPTMEWTXDKG 3362
Qy 358 -----VWVLNDEAGMOCCLSDSCQVILS-----NIKVLPTW 390
Db 3363 ELECTGKLEIKIADSFYILINDSSR-----RDSGAYILTADPPGFAKHIFNVVLDR- 3416
Qy 391 STPVPRASALPAPPTGSLPPOPTASALPDPAPASALPALAVISFLGLGLGVA 446
Db 3417 -----PGPEEG-----PLAVSEVTSEKCVLSWLPJDDGSA 3447

RESULT 39
LJMSNL
neural cell adhesion molecule 1 precursor, long domain splice form - mouse
N/Alternate names: NCAM-180
N/Contains: neural cell adhesion molecule, short domain splice form (NCAM-140)
C/Species: Mus musculus (house mouse)
C/Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text change 31-Dec-2000
C/Accession: A29673; S00844; S00384; A28281; A44290; S00383
R/Bartels, D.; Santoni, M.J.; Wille, W.; Rupprecht, C.; Chalk, J.C.; Hirsch, M.R.; Fontec
EMBO J. 6, 907-914, 1987
A>Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000
A/Reference number: A29673; MUID:87246524; PMID:3595563
A/Accession: A29673
A/Molecule type: mRNA
A/Residues: 1-548; 'T', 550-571, 'T', 573-574, 'D', 576-588, 'MQPS', 593, 'S', 595-599, 'P', 601, 'L'
A/Cross-references: EMBL:Y00051; NID:953342; PIDN:CAA68263.1; PID:953343
R/Santoni, M.J.; Bartels, D.; Barbats, J.A.; Hirsch, M.R.; Steinmetz, M.; Goridis, C.; W
Nucleic Acids Res. 15, 8621-8641, 1987
A>Title: Analysis of cDNA clones that code for the transmembrane forms of the mouse neu
A/Reference number: S00844; MUID:88067687; PMID:3684567
A/Accession: S00844
A/Molecule type: mRNA
A/Residues: 529-809, 1077-1115 <SN>
A/Cross-references: EMBL:X06328; NID:953322; PIDN:CAA29641.1; PID:9817984
R/Barbas, J.A.; Chalk, J.C.; Steinmetz, M.; Goridis, C.
EMBO J. 7, 625-632, 1988
A>Title: Differential splicing and alternative polyadenylation generates distinct NCAM E
A/Reference number: S00382; MUID:88283628; PMID:3396534
A/Accession: S00384
A/Molecule type: DNA
A/Residues: 642-1115 <BAR>
A/Cross-references: EMBL:X07195
R/Bartels, D.; Vopper, G.; Wille, W.
Nucleic Acids Res. 16, 4217-4225, 1988
A>Title: NCAM-180, the large isoform of the neural cell adhesion molecule of the mouse,
A/Reference number: A28281; MUID:88247737; PMID:2454455
A/Accession: A28281

```

```

A/Molecule type: mRNA
A/Residues: 804-1081 <BA3>
A/Cross-references: EMBL:X07244; NID:953321; PIDN:CAA30230.1; PID:g929720
R/Rougon, G.; Marshak, D.R.
J. Biol. Chem. 261, 3396-3401, 1986
A>Title: Structural and immunological characterization of the amino-terminal domain of m
A/Reference number: A44290; MUID:86140120; PMID:3512556
A/Accession: A44290
A/Molecule type: protein
A/Residues: 20-36 <ROU>
C/Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol
C/Comment: Several forms of NCAM are produced by alternative splicing. See also PIR:LJMS
C/Genetics:
A/Map position: 9
A/Intons: 643/3; 701/1; 770/2; 809/2; 1076/2
C/Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu
C/Keywords: alternative splicing; brain; cell adhesion; duplication; heparin binding; st
F/1-19/Domain: signal sequence #status predicted <SIG>
F/20-1115/Product: neural cell adhesion molecule, long domain splice form #status experi
F/20-809,1077-1115/Product: neural cell adhesion molecule, short domain splice form #sta
F/20-711/Domain: extracellular #status predicted <EXT>
F/34-98/Domain: immunoglobulin homology <IMM1>
F/132-191/Domain: immunoglobulin homology <IMM2>
F/152-156/Region: heparin binding #status predicted
F/161-165/Region: heparin binding #status predicted
F/228-290/Domain: immunoglobulin homology <IMM3>
F/262-272/Region: NCAM binding #status predicted
F/323-388/Domain: immunoglobulin homology <IMM4>
F/420-482/Domain: immunoglobulin homology <IMM5>
F/519-596/Domain: fibronectin type III repeat homology <FN3A>
F/625-665/Domain: fibronectin type III repeat homology <FN3B>
F/712-729/Domain: transmembrane #status predicted <TM>
F/730-1115/Domain: intracellular #status predicted <INT>
F/41-96,139-189,235-288,330-386,427-480/Disulfide bonds: #status predicted
F/222,316,348,424,450,479/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.3%; Score 124; DB 1; Length 1115;
Best local similarity 21.9%; Pred. No. 2.9;
Matches 68; Conservative 47; Mismatches 128; Indels 68; Gaps 15;

Qy 3 RG-VPRHLLVLDLALPAATQGNKVVLTCTASQKSIQFTM-KNSNQIK 60
Db 196 RGEINFDIOQIVAVPPTVQARQSIIVANATLGGSVTLVCDADPPEPTMTWGDGEIE 255
Qy 61 ILNQGSFLTYGSPKLNDRAD--SRSLWDQGNPPLIIXLKIEDSDTYIEVE---DQ 114
Db 256 -----NEEDERSRSVSDDSE--VTRNVDRKXDEAYVCIAENKAGEQ 297
Qy 115 KEVQLVFG---LTANSPTHLLQGOSLTLTLESPGSSPSVQCR-----SPRGKNIQ 164
Db 298 DASIHLLKVFAPKKTYYENQFAMELEEVYLTGASGDPISITWRSTRNISSEBQDL 357
Qy 165 GG-----KTSVSOLELQDSGTWTCYVLQONKVEFKIDIVLAFQKASSI----- 210
Db 358 GHMVRSHARVSSLTLSIQYRDAGEYCTASNTIGDQSGID---LEFQVAPLKQGPVA 414
Qy 211 VYKKEGQVE-----FSFPLAFVTEKLTGSGELMWQAEARASSSKSWITFDLKN-KEYSVK 264
Db 415 VYTWEGNQVNITCEVFAPSA-TIS-----WFRDQQLPSNYSNIKIYNTPSASYL 465
Qy 265 RVTQDPKLGKMG 275
Db 466 EVTPDSENDFG 476

RESULT 40
A64157
hypothetical protein HI0696 - Haemophilus influenzae (strain Rd KW20)
C/Species: Haemophilus influenzae
C/Date: 18-Aug-1995 #sequence revision 18-Aug-1995 #text change 29-Sep-1999
C/Accession: A64157
R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A

```



```

Db      58 EG-----KAPQRLLYDVASTARDVLESGSPGKYTHTRRRMGLRLQNLINDSGVYYC 113
Qy      110 EVEDQKEVOLVLEFULTANSDTHLLQGSGLT---LESPSSSSSVQCRSPRGNIOGG 166
Db      114 ATWDRPRLKLT--FG-----SGTTLVTTDQQLADVSPKPTIFLP----- 151
Qy      167 KTLSSQLELODGSMTCTVLQNKQKVEFKIDIVLAFC--KASIVYKKEGE----- 217
Db      152 ---SIAETKLOKAGTYLLCL-----EKFPPIIKIHQEKSNITLQSGSNTWKMTDT 202
Qy      218 QVESFPPLAFVTEKLTGSGELMWOAERASSSKSWITFDLKNKESVKEVETDPPKLOMGKK 277
Db      203 YMKFSFM-LTVPRESLDKEHRC--IVRHENNNKGIDGELIPEPITDVTDTYDPPKYNKSD 258
Qy      278 LPLHLTLOALPQVAGSGNLTALPAKTKGLHQEVNLYVMRTQLOKQLTC 328
Db      259 ANDVITM-DPKDMGSKDNDTLTLQTLTNTSAVYTYLLTLKSVVFAITTC 308

RESULT 43
ITCHNL
neural cell adhesion molecule long domain form precursor - chicken
N:Alternate names: NCM-180
N:Contains: neural cell adhesion molecule, short domain form (NCAM-140)
C:Species: Gallus gallus (chicken)
C:Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text change 22-Jun-1999
C:Accession: A43613; B43613; A25435; B25435; A46550; S36950; A46469; A60852; S29668
R:Cunningham, B.A.; Hemperly, J.C.; Murray, B.A.; Prediger, E.A.; Brackenbury, R.; Edelman
Science 236, 799-806, 1987
A:Title: Neural cell adhesion molecule: structure, immunoglobulin-like domains, cell sur
A:Accession: A43613; MUID:87206190; PMID:576199
A:Molecule type: mRNA
A:Residues: 1-175 <CU>
A:Cross-references: GB:M15860
A:Accession: B43613
A:Molecule type: protein
A:Residues: 20-44;120-127;202-221;320-342;399-415;640-659;822-828 <CUN>
A>Note: Asn-222 probably binds carbohydrate; Asn-226 probably does not
R:Hemperly, J.C.; Murray, B.A.; Edelman, G.M.; Cunningham, B.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 3037-3041, 1986
A:Title: Sequence of a cDNA clone encoding the polysialic acid-rich and cytoplasmic dome
A:Reference number: A25435; MUID:86206089; PMID:3458261
A:Accession: A25435
A:Molecule type: mRNA
A:Residues: 128-1091 <HEM>
A:Cross-references: GB:M13210
A:Accession: B25435
A:Molecule type: protein
A:Residues: 128-140;222-240;428-439;611-631;744-760;763-781;1080-1084 <HE2>
R:Murray, B.A.; Owens, G.C.; Prediger, E.A.; Crossin, K.L.; Cunningham, B.A.; Edelman, G
J. Cell Biol. 103, 1431-1439, 1986
A:Title: Cell surface modulation of the neural cell adhesion molecule resulting from al
A:Reference number: A46550; MUID:87033934; PMID:3771645
A:Accession: A46550
A:Molecule type: DNA
A:Residues: 810-1070 <MUR>
A:Cross-references: GB:X04479
R:Saenger, M.; Covault, J.
submitted to the EMBL Data Library, February 1993
A:Reference number: S36950
A:Accession: S36950
A:Molecule type: DNA
A:Residues: 1-17 <SAS>
A:Cross-references: EMBL:X70342; NID:G417631; PIDN:CAA49807.1; PID:G417632
R:Colwell, G.; Li, B.; Forrest, D.; Brackenbury, R.
Genomics 14, 875-882, 1992
A:Title: Conserved regulatory elements in the promoter region of the N-CAM gene.
A:Reference number: A44369; MUID:93122797; PMID:1478668
A:Accession: A44369
A:Molecule type: DNA
A:Residues: 1-17 <COF>
A:Cross-references: EMBL:Z12128; NID:G63653; PIDN:CAA78113.1; PID:G63654

```

```

A:Experimental source: White Leghorn
R:Colle, G.J.; Loewy, A.; Cross, N.V.; Akesson, R.; Glaeser, L.
J. Cell Biol. 103, 1739-1744, 1986
A:Title: Topographic localization of the heparin-binding domain of the neural cell adhes
A:Reference number: A60852; MUID:87057627; PMID:2430978
A:Accession: A60852
A:Molecule type: protein
A:Residues: 20-29 <COL>
R: Rao, Y.; Wu, X.F.; Gartezy, J.; Rutishauser, U.; Sin, C.H.
J. Cell Biol. 118, 937-949, 1992
A:Title: Identification of a peptide sequence involved in homophilic binding in the neur
A:Reference number: A43280; MUID:92363934; PMID:1380002
A:Contents: annotation: homophilic binding region
C:Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol
C:Superfamily: Various forms of NCAM are produced by alternative splicing.
C:Keywords: alternative splicing; brain; cell adhesion; duplication; heparin binding; si
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1091/Product: neural cell adhesion molecule, long domain form #status experimental
F:20-809;1071-1091/Product: neural cell adhesion molecule, short domain form #status exp
F:20-711/Domain: extracellular #status predicted <EXT>
F:34-98/Domain: immunoglobulin homology <IMM2>
F:132-191/Domain: immunoglobulin homology <IMM1>
F:152-156/Region: heparin binding #status predicted
F:161-165/Region: heparin binding #status predicted
F:228-289/Domain: immunoglobulin homology <IMM3>
F:262-271/Region: NCAM binding #status experimental
F:322-387/Domain: immunoglobulin homology <IMM4>
F:419-481/Domain: immunoglobulin homology <IMM5>
F:518-595/Domain: fibronectin type III repeat homology <FN3A>
F:624-665/Domain: fibronectin type III repeat homology <FN3B>
F:712-729/Domain: transmembrane #status predicted <TM>
F:730-1091/Domain: intracellular #status predicted <INT>
F:41-96;139-189;235-287;329-385;426-479/Disulfide bonds: #status predicted
F:222/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:315;347;423;449;478/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.3%; Score 123.5; DB 1; Length 1091;
Best Local Similarity 20.2%; Pred. No. 3;
Matches 85; Conservative 51; Mismatches 157; Indels 127; Gaps 16;

Qy      32 KKGPVELTCTAAGSKKSIQFMKNSNOIKILNGSGSFLTKGPKSLNBRADSRSLMDGN 91
Db      130 KEGDVAIVCDVSSLPPTITWKKGRDVLKQVRIV-----LSN 171
Qy      92 FPLIKKLIKEDSDTYICE-----VEDQKEVOLV-----FGLTANSDTHL 133
Db      172 NYLQIRGIKKTDEGTVCGRILARGEINFQIVIVNPPSVARQSTWMTATN----- 226
Qy      134 LQGSGLTITLESPPGSSPVQCRSPRKNIO-----GKTLSSQLELODSGT 181
Db      227 -LSGSVTLACDADPPEPTWTM--TKDGEPIEQEDNEKRSFNYDGSSEILIKVAKSDPAE 284
Qy      182 WTCVVLQNKQKVEFKIDIVLAFCQKASSIYKKEGDEVESFPPLAFVTEKLTGSGE----- 237
Db      285 YICIAENKAGQDDTIHLKYPAPK---ITYVEKTMLELDQITLICE---ASGDPIPS 338
Qy      238 LMWOAERASSSKSWITFDLKNKESVVRVTDPPKLOMGKKLPLHLTLOALPQVAGSGNT 297
Db      339 ITWTSFRNISNEKKTIDGRIVRSHARVSS-----LTIKEL--QYTDAGEY 383
Qy      298 TLAEAKTKGLHQEVNLYVMRTQLO-----QKULTCVWGPTSKMLSLKLE 345
Db      384 VCTASNTIGQDSQAMYLEVOYAPPLQGPVAVTYTEGQNVITCEVF--AYPSAIVS---- 437
Qy      346 NKEAKVSKREKPVVVLNPEAGMQLCLSDSGQVLLSEINIKVLPWTWSPVPRASALPAP 405
Db      438 -----WFRD-----GQLPSSNYSNIKITYNT---PSASYLEVTP 468

RESULT 44
A38096
perlecan precursor - human

```

[illegible][illegible]

Best Local Similarity 18.7%, Pred. No. 4.9; Mismatches 153; Indels 222; Gaps 22;
Matches 100; Conservative 59;

Dy 32 KKGDVETLTCTASOKSIQIFHWKNSNOIKLIGNOSFLYTKGPSKLNDRADRSRLMDQGN 91
Db 99 EKGNKLLIECTYTGTGPPISTVMKKNGVILKHSKCSITTTETA----- 143
Oy 92 FPLIINKLKIDSDTYTICEVEDQKEEVLVFGLTANSDPHLLLOGSITLTLESPP---- 147
Db 144 --LEIPNPKLEDQOCQSCHIEND-----SGDONCH-----GATILEPPEFVT 184
Oy 148 -----SSPSVQCSRPRGNIOGGKTLVSQGLELDGSGTWTCYTLNQKKVEEKID 198
Db 185 PLEVQVTVGDSASIQCQ-----VAGTPMIVSYKGYDTLRGATV-----KHFRNQ 233
Oy 199 IVLVIAFOKASSIVYKKEGEQVEFPFLAFYTEUKLTGSGLMMQAERASS----- 248
Db 234 VATLVFSQVDS-----DDSGEYICXVENTYTGENTSSLLTVGRKLPSPSTRKLRDVHE 287
Oy 249 -----KSWITFDLNKE-----VSVKVTDQPKLQWKKL 278
Db 288 TVGLPVTFPDCCIAGESEPIEVSMFKDNVAKEDYNHTSFIDNVAILQLIKTKDKSLMQ-- 345
Oy 279 PLHLTLPALQVYAGSGULTLALEKT----- 305
Db 346 -YTCASNAGITASSSGKLVLT-EGKTPPFEDPTPTPYDGIIGESADFECHISGTQPIRV 403
Oy 306 --GKLHQV-----NLVVRATOLOK-NLTC-----EWMGPTSPMLSLK 343
Db 404 TWAQDNGEIRTCGVNYQISYVENTAHLTLIRDRDGSKRTCYASNEV-GKDSCTAQLANVK 462
Oy 344 -----LENKEAKVSKR---EKPV---WVLN----- 362
Db 463 ERKTPPTPTRKLSEAVEETEENELKEGRVAVAGSQLFTVSWYKNQNEBVHSPCEISPFN 522
Oy 363 -----PEAGMOCCLLS-DSGCVLTLESNT-----KYLPTWSTPVMHRAS 399
Db 523 TLHLHKSVGQSDAGLYTCRKVNSEGAYLCTSSVVIREFPKRPVPDFLOPAAT 576

RESULT 46
JC4776
limbic-system-associated membrane protein precursor - human
C:Species: Homo sapiens (man)
C:Date: 10-May-1996 #sequence_revision 16-Aug-1996 #text_change 19-May-2000
C:Accession: JC4776
R:Pimenta, A.F.; Fletcher, I.; Levitt, P.
Gene 170, 189-195, 1996
A:Title: cDNA cloning and structural analysis of the human limbic-system-associated mem-
A:Reference number: JC4776; MUID:96235133; PMID:8666243
A:Accession: JC4776
A:Molecule type: mRNA
A:Residues: 1-338 <PM>
A:Cross-references: GS:U41901; NID:g1276898; PID:NAC50569.1; PID:g1276899
A:Experimental source: brain
A:Comment: This is a neuronal surface glycoprotein distributed in cortical and subconfid-
C:Genetics:
A:Gene: lamp
C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termimn
C:Keywords: brain; glycoprotein; membrane protein; phosphoprotein
F:1-7/Domain: signal sequence #status predicted <Sto>
F:333-338/Region: hydrophobic
F:46:66:136:148:279:287:300:315/Binding site: carbohydrate (asn) (covalent) #status preded
F:42:115:142:154:171:220:231/Binding site: phosphate (thr) (covalent) #status predicted
F:95:192:204:236:310/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 5.2%; Score 121.5; DB 2; Length 338;
Best Local Similarity 20.7%; Pred. No. 0.85;
Matches 86; Conservative 63; Mismatches 123; Indels 143; Gaps 22;

Oy 10 LLLVLIQLALLPA-----TQGNKVLAGKKDVELTCTASQKSIQIFHWKNSNOIKI 61
Db 14 LVLLRIQLLTGLGVPVRSDVDNRGTDNITVGCGDAIRCVLEBKNS-KVAMLRSGIIIF 72

Oy 62 LGNGSFLTKGSKLNDRA--DSRSLMDQGNFPLIIINLKIDSDPYIGVEVDQKE--- 116
 Db 73 AGHD-----KMSLDPRVLEKEHRSL-----EYSLSLIQKVDVDEGYSQVQTOHEPPT 121
 Oy 117 -EVOLLVFG----LTPNSDPTLLQGSGLTFLTLESPPGSSPSVOCR--SPRGKNTQGGKT- 168
 Db 122 SQVLIIVQVPPKISNISSDVYTNEGSNVTLVCMANGRPEPVITWRHLTPTGREGEGEEX 181
 Oy 169 LSVSGLBJODSGTWTCTVYLQNGKQVEFKIDIVLAFQKASSIYKKEGEQVE--FSFPLA 226
 Db 182 LEILGITREQSGKYEC-----KANENSSADVKKQVTVNYPPT 220
 Oy 227 FTVEKLTGSGEL-----WMQAEPRASSSKSWITFDKXNE----V 261
 Db 221 ITESK---SNEATTRQASLKCBCASAVPAPDEFEMWDDTRINSANGV--ETKSTEGQSSL 275
 Oy 262 SVKRTYODPKLQMGKKLPLHLTLPOLPYAGSGNTLLALEATGKGLHOEVNLYVMEATQ 321
 Db 276 TVTNTTEE-----HY---GNTTCVANAKLGVTN--ASLVLFR--- 307
 Oy 322 LQKNTLCEWGGTSPKLM---LSLKLENKBAKYSKREKPVWVLNPEAGMOCILIS 373
 Db 308 -----PGSVRGINGSISLAV-----PLMLL---AASLILCLIS 336

RESULT 47
 S28061
 SCPI protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 08-Oct-1999
 C:Accession: S28061
 R:Miscissem, R.L.U.; Offenberg, H.H.; Dietrich, A.J.U.; Riesewijk, A.; van Iersel, M.; H
 EMBL J. 11, 5091-5100, 1992
 A:Title: A coiled-coil related protein specific for synapsed regions of meiotic prophase
 A:Reference number: S28061; MUID:93099884; PMID:1464329
 A:Accession: S28061
 A:Molecule type: mRNA
 A:Residues: 1-946 <MEU>
 A:Cross-references: EMBL:X67805; NID:g57212; PIDN:CAA48006.1; PID:g57213
 C:Genetics:
 A:Gene: SCPI
 C:Keywords: DNA binding

Query Match 5.2%; Score 121; DB 2; Length 946;
 Best Local Similarity 19.8%; Pred. No.3.7; Mismatches 118; Gaps 24;
 Matches 96; Conservativity 95;

Oy 9 HLLVLTLOL----ALPPAATQGNKVVGLGKGDVTELCTASQKSQISQPH---WKNNSQIK 60
 Db 323 HSLVVELKATTCCTLEBELLRTGQRLNENDDQKLITMELQKSSSELEMTKPFKNKKEVE 382
 Oy 61 ILNGGSEFLTKGSPSKLNDR-----ADSRSLMDQGNFPLIIINLKIED----- 103
 Db 383 -LEELKTLIAEDQKLLDEKKQVEKLAEBLQSGEQSLFTLLQTRKEKHIDEVQVTVTKTS 441
 Oy 104 SDTYICEVEDQKEVQV---LLVFGLTANSPTDTHLLOQSGLT-----LTLSEPPGSSPSVQC 155
 Db 442 BEHYIKQVEEMKTELEKEKLNIELTANSDMILLKKNKLVQEASDMVLELKQHQEDITIC 501
 Oy 156 RSPRGKNTQGGKTLVSQLELQDS--GTWTCVYLQNGKQVEFKIDIVLAFQKASSIYV-- 212
 Db 502 KKQGERMLKQIETLEKEMNLNDELSEVAKREFIQGDGEVQCKD---KSBNANASIEIEV 558
 Oy 213 -KKEGE-----QVEFSFPLAFTVEKLTGSGELWMQAEPRASSSKSWITFDLK- 257
 Db 559 LKKEKQMKLEBKCNMLKKQIEENK--SKNIEBELHOENKA-LKKSSAEAKQQLNAYEIKV 614
 Oy 258 -----NKEVSVKRTYODPKLQMGKKLPLHLTLPOLPYAGSGN 296
 Db 615 NKLELELASTKQKFEEMINNYQKEIEIKKISE-KV-LGEVEKAKAVADEIKV---QKE 669
 Oy 297 LTLALEATGKGL-----HOEVNLVWRATD--LQKNTLCEWGGTSPKMLSLKLEN 346

```

Db      670 IDLCOHKIAEMVVALMEKHQDYKIVBERDSEGLYKNRROE---OSSAKVALETLSIN 726
Qy      347 -----KEAKYKREKPVWVNLNPEAGMMQCLSDSGVLLSNIKVLP---TW----- 390
Db      727 IRNELVSLKKOLEVEKEKEKXKNRQEN---TAILDKKKKIQASLSESPKATSMKFD 783
Qy      391 -STP 393
Db      784 KTFP 787

```

```

RESULT 48
151669
tumor suppressor - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C/Accession: F51669
R/Pierceli, W.E.; Reale, M.A.; Candia, A.F.; Wright, C.V.; Cho, K.R.; Fearon, E.R.
Dev. Biol. 166, 654-665, 1994
A/Title: Expression of a homologue of the deleted in colorectal cancer (DCC) gene in the
A/Reference number: F51668; MUID:95113183; PMID:7813784
A/Accession: F51669
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1427 <PIE>
A/Cross-references: EMBL:U10986; NID:g606873; PIDN:AAA70168.1; PID:g606874
C/Genetics:
A/Gene: XDCCA

```

```

Query Match      5.2%; Score 121; DB 2; Length 1427;
Best Local Similarity 20.8%; Pred. No. 6.4;
Matches 94; Conservative 64; Mismatches 160; Indels 126; Gaps 24;

Qy      34 GDTVELTCTASQKSIQFHW-KNSNQIKILNQSGFLTKGPKSLNDRADSRSLMDGDNF 92
Db      154 GDTALLREITGEPMPTISWQNEEDLKV-----TPG-----DPRLLVLPST- 196
Qy      93 PLIKNLKIEDSDTYICEVD-----QKEEVQL-----VFGLTANSDTHLLOG 136
Db      197 -LQISRLQTAGGYRCLAKNGSARVGENAEALRLSSGLHQQVF-LQRPNVVAIEG 254
Qy      137 QSLTL-----TLSPGSSPSVOCSPRGKNIOGKTLVSQLELDSDGTWCTV 186
Db      255 QDAVLECAVSGYPTPIVMMQGDDEP-VPIRT-RKSYLGGSNLLISNVTDDAAYTCVA 312
Qy      187 LQNKKVEFKIDIVLA---FQKASIIYKKEGQVESFPLA---FTVEKLTGSGEL-- 238
Db      313 TYKNENISFSADLTVMWPPQFLNHPANLYAYESMDIEFECVSGKPSFTVMTXNGEYVI 372
Qy      239 ---MWQAEKAS-----SSKSWITPDLNKEVSVKRVTO---DPKIQMGKKLPVHL 282
Db      373 PSDYFOIYDGSNLRILGLVKSDEGYOCIAENAGNIOYVQQLIIPDAVSSSILP--- 429
Qy      283 TLPOAL-----PQYAGSGNL---TLAEAKGKLHQEVNLVVMRATQIQ- 323
Db      430 SAPRVVVLVVSREVRLSMRPPVESKGNIOITYYFSGQVQREBRAVNTSQPISTLQITV 489
Qy      324 KNLICE-----VMGP-----TSPKLMSLLENKE-----AKVSRRE 355
Db      490 GNLTPFEETVYRNVVAYNMKGEGSSQEVKVVYQFLEQVPEVNDQVVSTPTSVLSWD 549
Qy      356 KPVVNLNPEAG-MWQCLSDSGVLLSNINIKV 386
Db      550 PPAVANGVGVGRLECAETFSGR---EQNIEV 578

```

```

RESULT 49
A31923
amalgam protein precursor - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 21-Jul-2000
C/Accession: A31923

```

```

R/Seeger, M.A.; Hatfield, L.; Kaufman, T.C.
Cell 55, 589-600, 1988
A/Title: Characterization of amalgam: a member of the immunoglobulin superfamily from Drc
A/Reference number: A31923; MUID:89028670; PMID:3141062
A/Accession: A31923
A/Molecule type: DNA
A/Residues: 1-333 <SEE>
A/Cross-references: GB:M23561; NID:g156920; PIDN:AAA28367.1; PID:g156921
C/Genetics:
A/Gene: FLYBase:Ama
A/Cross-references: FLYBase:FBgn0000071

```

```

Query Match      5.2%; Score 120.5; DB 2; Length 333;
Best Local Similarity 19.4%; Pred. No. 0.97;
Matches 56; Conservative 48; Mismatches 123; Indels 61; Gaps 9;

Qy      10 LLVLQIAL-----LPAATGKNVVLGKGDVTELTCTASQKSIQFHW-----KNSNQ 58
Db      10 LIFCLALSDLSVLSAPVLSQISKDVVASVGSVEFNCVTEVEVQLSVSMKRPSESDTNS 69
Qy      59 I-----KILG-----NQSFITKGPSPKLNDRADSRSLMDGDNFPLIKNLKIEDSDTYIC 109
Db      70 VLISMKNILSLPKRKNVTVTEGPKT-----GSAITFRQNIIEVDMGEYEC 117
Qy      110 EV-----EDQKEVQLT---VFGLTANSDTHLLOGSITLTLESPPGSSPSV----- 153
Db      118 QVAVSATSEKTKSLSIQIKTPPIAENTPKSTLVTEGQNLDELTCHANGFPKPTISWAREH 177
Qy      154 QCRSPRGKNIOGKTLVSQLELDSDGTWCTVYLQNKVEFKIDIVLAFOKASIVYK 213
Db      178 NAWMPAGGHILAEPTLIRSVHRMDRGYYCIAQNGEGQDCKLIRVEVFRPOIAVQRP 237
Qy      214 KEQGEVFEFSFPLAFTVEKLTGSGELWQ-----AERASSS 248
Db      238 KIQMWSHSALELCSVQGYAPFVWVHKNGVPLQSSRHHEVANTASSS 285

```

```

RESULT 50
ORRG
secretory component precursor - rabbit
N/Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor
N/Contains: free secretory component; transmembrane secretory component
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 16-Jul-1999
C/Accession: A02111; A28077
R/Moskov, K.E.; Friedlander, M.; Blobel, G.
Nature 308, 37-43, 1984
A/Title: The receptor for trans epithelial transport of IgA and IgM contains multiple imm
A/Reference number: A02111; MUID:84142246; PMID:6322002
A/Accession: A02111
A/Molecule type: mRNA
A/Residues: 1-773 <MOS>
A/Note: the authors translated the codon ACC for residue 54 as Asn
R/Fructiger, S.; Hughes, G.O.; Hanly, W.C.; Jaton, J.C.
J. Biol. Chem. 263, 8120-8125, 1988
A/Title: Rabbit secretory components of different allotypes vary in their carbohydrate co
A/Reference number: A28077; MUID:88228032; PMID:3131339
A/Accession: A28077
A/Molecule type: protein
A/Residues: 87-114;410-424 <FRU>
C/Comment: This receptor binds polymeric IgA and IgM at the basolateral surface of epithe
process, cleavage occurs to separate the extracellular portion, also known as the secret
C/Comment: The five domains exhibit homology with immunoglobulin V regions. The similarit
C/Comment: Alternative splicing in the extracellular domain leads to high or low molecule
C/Superfamily: secretory component; immunoglobulin homology
C/Keywords: alternative splicing; duplication; glycoprotein; immunoglobulin receptor; pol
F/1-18/Domain: signal sequence #status predicted <SIG>
F/19-773/Product: transmembrane secretory component #status predicted <MATW>
F/30-647/Domain: extracellular #status predicted <EXT>
F/39-117/Domain: immunoglobulin homology <IML>
F/148-227/Domain: immunoglobulin homology <IM2>

```

F:253-326/Domain: immunoglobulin homology <IM3>
 F:362-440/Domain: immunoglobulin homology <IM4>
 F:471-540/Domain: immunoglobulin homology <IM5>
 F:648-670/Domain: transmembrane #status predicted <TM>
 F:671-777/Domain: intracellular #status predicted <INT>
 F:46-115,155-223,369-438,478-538/Disulfide Bonds: #status predicted (partial) #status experimental
 F:108/Binding site: carbohydrate (Aen) (covalent) #status experimental
 F:418/Binding site: carbohydrate (Aen) (covalent) #status experimental

Query Match 5.2%; Score 120.5; DB 1; Length 773;

Best Local Similarity 20.3%; Pred. No. 3;

Matches 108; Conservative 75; Mismatches 199; Indels 149; Gaps 26;

```

QY 5 VEFRIHL-----VLQALLPAAQGN-----KVLGKKGDTVELTCAQSKKQ 50
DB 210 VTIKILQNDAGQVYQSGSDPTAEQNVDLRLPGILYGLGGSVTFECALDSEDANA 269
QY 51 FHMKNQK-----ILNGSFLTKGPSKLANDRADSRSLW---DQGNFPLIKKLTIED 103
DB 270 V--ASLRQVRGGNVVIDSGTI-----DPAEGRLFPKAEHGHSVVIAGLRKED 318
QY 104 SPTTYICEVEDQKEVQLVFGLTANSDFHLQ---GQSILTLLEBP-----FGSSPVQC 155
DB 319 TGNVLCGVQSN-----GQSGDGPQLRLQFVNEBIDVSRSPVLKGFPGSGVTRIC 369
QY 156 -RSPR-----GKNL-----GGKTLV--SOLE 175
DB 370 PYNPKRSDSHLOLYMEGQTRHLVDSGEGVQKDYGRLLAFEPENGITSVLANQRT 429
QY 176 LQDSGTWTCTVLQNKQKVEFKIDIVLAFQKASSI--VYKKEGQVRS--FPLAFTEK 231
DB 430 AEDEGFYWC-VGDDDESILTSVKLQIVGDESPPTIDKFTAGGEVETICHPCRY---- 484
QY 232 LTGSELWQAEARASSSKSWITF-DLKKEVSVKVTDQPKLQMGKULPHLTLPQALPQ 290
DB 485 -----FSEKTKCMNKHGCCEDLPTKLSSSGDLVKCNMNVLTLLTLDVSED 531
QY 291 YAG---SGNLTLALEAKTGKHLQHEVNLVVMRATOLQKULTCGVWPTSPKMLSLKLNK 347
DB 532 DEGWYWC-----AKDQHEFEVAAVAVELTEPAKVAVEPAKYPVDAKAAAPPAEAK 584
QY 348 EAKVSKREKPVVNLPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVHPRAALPAPPTG 407
DB 585 ----AKACFPV-----PRRQWYPL-----SRKLRT-SCP-ERLLAEVAAVQS 622
QY 408 SALPPQTRASALPDPPAASALPALAVISFLGLGL-----GVACVLARTR 453
DB 623 AEDPASGSRASVDASSAGSGSAKVLISTVLPLGLVLAAGAMAVAIARAR 673

```

RESULT 51

T18358
 apolipophorin precursor protein - tobacco hornworm
 C/Species: Manduca sexta (tobacco hornworm)
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
 C/Accession: T18358
 R:Sundermeyer, K.; Hendricks, J.K.; Praead, S.V.; Wells, M.A.
 Insect Biochem. Mol. Biol. 26, 735-738, 1996
 A>Title: The precursor protein of the structural apolipophorins of lipophorin: cDNA and
 A/Accession: T18358
 A/Reference number: Z18891; MUID:97166603; PMID:9014323
 A:Structure: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-3305 <SUN>
 A:Cross-References: EMBL:U57651; NID:g1399217; PID:g1399218; PIDN:AAB53254.1

Query Match 5.2%; Score 120.5; DB 2; Length 3305;

Best Local Similarity 22.5%; Pred. No. 21;

Matches 97; Conservative 66; Mismatches 176; Indels 93; Gaps 21;

```

QY 26 NKVVLGKKGDVLTCTASQKKS-----IQFMKNSNOIKILNGSGFLTKGPSKLANDRA 80
DB 1852 NSVVVDADGRVYKIDSSIVLSKAHPVLDIQVHSPSSDKIRRLYLQSSLSSTQGLTEKVV 1911

```

```

QY 81 DSRRLMDQGNFPLIKKLTIEDDTYICEVEDQKEVQLVFGLTANSDFHLQGSILT 140
DB 1912 DN-----INDICLD---AVSEANVQKDNV---AFKVVANAKELGMKNNGID 1951
QY 141 LTLESPPSSPSVQCRSPR-SKNIQGKTLTSVQLELQDS---GTWCTVLQNKQKVEFK 196
DB 1952 IS-SKDSGGRLEPHANNDKNVLSGSTSISNGEQKTLTIESGSKYKVEEQKSNFX 2010
QY 197 IDIVVLAFAKASSIVYKKEGQ-VESFPPLAFVTEKLTGSELWQAEARASSSKSWIT-F 254
DB 2011 Y-----IRVFTDSNKGVEYTFEPPVAL-----GERSYVAESRVTVNY 2046
QY 255 DLKKEVSVKVTDQPKLQMGKULPHLTLPQALPQVAGSNLTALAEAKTGKHLQHEVNL 314
DB 2047 EYKNSYVCEKKKCAHAIEQSK--IDWSTGMVNVNINAG-LDL--RKLGV----- 2093
QY 315 VVMRATOLQKULTCGVWPTSPKMLSLKLNKAEKVKREKPVVNL-PEAGMOCCLLS 373
DB 2094 ----APELGQMDREVDSDRRPRPTLDLHT-NKE-----DRKYHLAAVYTPENGHY----A 2140
QY 374 DSGOVLLESNIKVLPTWSTPVHPRAALPAPPTGSAALPD-----PQTRASALPDPPA 425
DB 2141 SGVTWRLPSSRYMALE--YTLTHPTSDLPFPKGEACIDLDKNRPGHKTSAFLVDYSNS 2198
QY 426 SALPALAVISF 437
DB 2199 GSEDKAAVEIGF 2210

```

RESULT 52

hypothetical protein DKFZps6411922.1 - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 23-Jul-1999
 C/Accession: T08678
 R:Mambuti, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, May 1999
 A:Reference number: Z16469
 A:Accession: T08678
 A:Molecule type: mRNA
 A:Residues: 1-584 <MAN>
 A:Cross-References: EMBL:AL049946
 A:Experimental source: fetal brain; clone DKFZps6411922
 C/Genetics:
 A>Note: DKFZps6411922.1

Query Match 5.1%; Score 119; DB 2; Length 584;

Best Local Similarity 20.3%; Pred. No. 2.6;

Matches 69; Conservative 57; Mismatches 130; Indels 84; Gaps 17;

```

QY 12 LVTLQALLPAAQGN-----NKVVLGKKGDVLTCTASQKKSIOFMKNSNOIKILNQ 65
DB 285 LILQVTLVEPMEKPIFDHPISKITAMAGHTISLNCAGTPPSLVW-----VLPN- 336
QY 66 GSFLTKRP--SKANDRADSRSLMDQGNFPLIKKLTIEDDTYICEVEDQKEVQLV- 122
DB 337 GTDILQSGOQLQRFYHKADG-----LHISGLSSVDAAGACVARNNAAGHTERLVS 386
QY 123 --FGL--TANSDTG-----LLQGSILTLLESPPSSPSVQCRSPRKNIOGK----- 167
DB 387 LKVLKKEPASQYNNLVSIINGETLKPCTPRPAGQGRFSTLPLNGHNLBEPQLGKVS 446
QY 168 ---TLVSVQLELQDSGTWTCTVLQNKQKVEFKIDIVLAF-----QKASSIVYKKEGQ 218
DB 447 LDNGTLTVREASVDRGTIVYCRMETEGSPVTSIPVIVIAVPPRITSEPPVITRRGNT 506
QY 219 VEFSPPLAFVTEKLTGSELWQAEARASSSKSWITPDLKKEVSVKVTDQPKLQMGKUL 278
DB 507 VKLNC-MAMGIPK---ADITWELPDKSHLAAQVQARL-----YGNRF 544
QY 279 PLH-----LTLQALPQVAG-----SGNLTLALEAKTGKHL 309

```

Db 545 -LHPGSLTIQATQRDAGFYCKMAKNI-LGSDSKTYIYH 582

RESULT 53

JC4917

Signal transducing adaptor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 26-Sep-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jul-2000

C:Accession: JC4917

R:Takeishi, T.; Ariga, T.; Asao, H.; Tanaka, N.; Higuchi, M.; Kuroda, H.; Kaneko, K.; M

Biochem. Biophys. Res. Commun. 225, 1035-1039, 1996

A:Title: Cloning of a novel signal-transducing adaptor molecule containing an SH3 domain

A:Reference number: JC4916; MUID:96374438; PMID:8780729

A:Accession: JC4917

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-548 <TAK>

A:Cross-references: GB:U43900; NID:G1556460; PIDN:AA52840.1; PID:G3645912

A:Experimental source: T cell

C:Comment: This protein contains a Src-homology 3 domain and the immunoreceptor tyrosine

C:Genetics:

A:Map position: 2A2-B

C:Superfamily: SH3 homology

F:2-540/Product: signal transducing adaptor molecule #status predicted <MAT>

F:217-264/Domain: SH3 homology <SH3>

F:359-387/Region: immunoreceptor tyrosine-based activation motif

Query Match 5.1%; Score 118; DB 2; Length 548;

Best Local Similarity 21.0%; Pred. No. 2.8; Indels 110; Gaps 19;

Matches 92; Conservative 62; Mismatches 174; Indels 110; Gaps 19;

69 LTKGPKSLNDRADSRSLM--DOGNF-----LTKKLTIEDSDTYIC--EVEDQKEE 117

Db 97 LNKHPKCEKCKLALMVTEDPKNDPQLSLSAMIKLK--EGVTPFAISQAAEQAKA 155

Qy 118 VQLVLF--GLTNSDTHLQGSITLTLESPPGSSPSVQCRSPRGKNI-----QGSKT 168

Db 156 SPALVAKDPGVATVATKEEDLAKAIELSKQOQSAFVSTLYPSTSLTLTNHQEGKRV 215

Qy 169 LSVGLSELDGSGTCTVLQOKKVEFKIDIVLAFQASSIVYKKEEQVEFSPLAFT 228

Db 216 RAYVDFAEDNELT-----FRAGEITVLDDSDPNMKGETHQVGLFPSNFV 264

Qy 229 VEKLTGSGELMWQAEARASSKSMITFDLNKEVSKVKTQDPK-----LWQKG 276

Db 265 TADULTA-----EPERMITEKKTQVF--NDVQIETIEPEPEPAFTIDEDKMDQLQW-- 313

Qy 277 KLPLHLT-----LQALPQVAGSGNLTLAEKTKLHQ-----EVNLVVMRATQL 322

Db 314 --LQSTDPDNPQPLPELHLAEWCQMGPLIDKLDIDRKISELSELNVKVMELSL 370

Qy 323 QKNLTCEWGPISPKLMLSLKLENKEAVSKREKVVWLNPEAGMWQGLSDSQVLLS 382

Db 371 YTKLMNE--DPMYSMAKLGQSQVYLGSSAVASGVYRGP-----AQSGTYLVAG 418

Qy 383 NIKV--LPTWSTPVPRAS-----ALPAPPTGSLAPDPPTASALPD----- 421

Db 419 SAQWHLDSYSLPPEQLSSISOGAVPS--SANGALPQOOTQASYPNAAVSVQGSNYSRQA 477

Qy 422 ---PPAASALPALAVT 435

Db 478 SIYSPPAAMAAAAAAYV 495

RESULT 54

T29757

Protein UNC-89 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Dec-1999

C:Accession: T29757

R:Du, Z.; Le, T.T.; Wilson, R.

Submitted to the EMBL Data Library, May 1997

A:Description: The sequence of C. elegans cosmid C09D1.

A:Reference number: Z20679

A:Accession: T29757

A:Status: preliminary; translated from GB/EMBL/DB

A:Molecule type: DNA

A:Residues: 1-664 <DU2>

A:Cross-references: EMBL:AF003131, PIDN:AA854132.1; GSPDB:GNO00019; CESP:unc-89

A:Experimental source: strain Bristol N2; clone C09D1

C:Genetics:

A:Gene: CESP:unc-89

A:Map position: 1

A:Initons: 17/2; 108/3; 154/2; 211/2; 265/3; 326/2; 352/3; 426/2; 454/1; 500/1; 537/1; 63

/3; 5917/1; 6027/1; 6061/3; 6153/2; 6515/1; 6552/3; 6609/1

Query Match 5.1%; Score 118; DB 2; Length 6642;

Best Local Similarity 18.4%; Pred. No. 80; Indels 206; Gaps 22;

Matches 94; Conservative 67; Mismatches 145; Indels 206; Gaps 22;

Qy 32 KKGDVLELTCTASQKSIQFHW-KNSNQIKILGNQGSFLTGPKSLNDRADSRSLMDQG 90

Db 2899 KKGHSVAFECVVPDPTGVCCKMLKDGEIELIARIVQRTGP-----EG 2943

Qy 91 NF--PLIKKLTIEDSDTYICEV-----DQKE-----VQLLVFG 124

Db 2944 HITQELVLDVTPEDAGKTCIVENTAGKDTCEATLVLTLSLKSEKKAPEFVLQDX 3003

Qy 125 LTANSDDTHLQGSGL-----TLFLESPPGSSPSVQCRSPRGKNIQGGKTLVS 172

Db 3004 TTTSKENVLECKVIGEPKPKVSMHDNVSREKPNSEKTIQTESIVSEVEGERTIT 3063

Qy 173 QLELDGSGTCTVLQOKKVEFKIDIVLAFQASSIVYKKEEQVEFSPLAFTVEKL 232

Db 3064 SSELSHQKGYTC-----IAENTGTSKTEAF-----L 3090

Qy 233 TGSSELIMWQAEARASSKSMITFDLNKEVSKVKTQDPKQKMLPLHLTL--PQALP 289

Db 3091 TVQGE-----APVTEKELQNKELSI-----QBKVLSCSVGSPQPHV 3128

Qy 290 QVAGSGNLT-----LAEAKTGKLEQVN--LVWMRATQ----- 321

Db 3129 DFVFSFTTIVERTKITSSKIALE-----HDQTNHMRVVISQITIEDIVSKATATNS 3182

Qy 322 ---LQKNLTCEWGP-----TSPKLMLSLKENK-----EAKYSKREKV-- 358

Db 3183 IGTATSTSKITTKVPAVFPQGLKTSVKEKEIKNEVKVGSAPVDFPKDPVGEDG 3242

Qy 359 ---WVNLPEAGM-----QCLSDSQVLLSENIKVLPTSTVHPRAGSLPAPPTGSL 410

Db 3243 NHEMKKNPEGVFTLVVQKATTDAGKYTAKA-----SNPAGTASSAEAEVTS-L 3293

Qy 411 PDPQ-----TASALPDP 422

Db 3294 EKPTFVARELVTEVKINETAITLSVTKGVDP 3325

RESULT 55

IURTNC

neural cell adhesion molecule short domain form precursor - rat

N:Alternate names: NCAM-140

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999

C:Accession: S00846; B37795; I58136

R:Small, S.J.; Shull, G.E.; Santoni, M.J.; Akeson, R.

J. Cell Biol. 105, 2335-2345, 1987

A:Title: Identification of a cDNA clone that contains the complete coding sequence for a

A:Reference number: S00846; MUID:88059265; PMID:3680385

A:Accession: S00846

A:Molecule type: mRNA

A:Residues: 1-858 <SWA>

A:Cross-references: EMBL:X06564

R:Small, S.J.; Akeson, R.

J. Cell Biol. 111, 2089-2096, 1990

A:Title: Expression of the unique NCAM VASE exon is independently regulated in distinct

A:Reference number: A37795; MUID:91035620; PMID:1699951
 A:Accession: B37795
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 340-381 <SM2>
 R:Small, S.J.; Haines, S.L.; Akesson, R.A.
 Neuron 1, 1007-1017, 1988
 A:Title: Polypeptide variation in an N-CAM extracellular immunoglobulin-like fold is dev
 A:Reference number: 158136; MUID:90166485; PMID:2483093
 A:Accession: 158136
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 355-364 <RES>
 A:Cross-references: GB:M32611; NID:9205643; PIDN:AAA1679.1; PID:9205644
 C:Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol
 C:Comment: Various forms of NCAM are produced by alternative splicing.
 C:Gene: NCAM
 C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu
 C:Keywords: alternative splicing; brain; cell adhesion; duplication; heparin binding; si
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-858/Product: neural cell adhesion molecule, short domain form #status predicted <MA
 F:20-721/Domain: extracellular #status predicted <EXT>
 F:34-98/Domain: immunoglobulin homology <IMM1>
 F:132-191/Domain: immunoglobulin homology <IMM2>
 F:152-156/Region: heparin binding #status predicted
 F:161-165/Region: heparin binding #status predicted
 F:228-290/Domain: immunoglobulin homology <IMM3>
 F:263-272/Region: NCAM binding #status predicted
 F:323-398/Domain: immunoglobulin homology <IMM4>
 F:430-492/Domain: immunoglobulin homology <IMM5>
 F:529-606/Domain: fibronectin type III repeat homology <FN3A>
 F:635-698/Domain: fibronectin type III repeat homology <FN3B>
 F:722-733/Domain: intracellular #status predicted <INT>
 F:740-858/Domain: intracellular #status predicted <INT>
 F:41-96,139-189,235-288,330-396,437-490/Diulfide bonds: #status predicted
 F:222,316,348,434,460,489/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 5.1%; Score 117.5; DB 1; Length 858;
 Best Local Similarity 19.1%; Pred. No. 5.5;
 Matches 79; Conservative 52; Mismatches 143; Indels 139; Gaps 17;

3 RG-VEFRLHVLQALPAATGKNNKVLGKGGDTVELCTASQKKSIOFHH-KNSNOIK 60
 196 RGEINFKDIQVTVNVPYQARQSIYNATANGSVTLVCDAGPEPTMSTKDEEIE 255
 61 ILNGSSFLTKGPRSLNDRADRSRLMDGNFPLIKLIKIEDSDTYICEV---DQKE 116
 256 -----NEEEDDEKHIIFSDSSELTITRVNDKNDDEAYCIENKAGEDDA 299
 117 EVQLLVFG---LTANSDTHLQGSGLTTLESPPGSSPSVOCRRPRGKNIOGKTLVS 172
 300 SIHLVVFAPKPIYVENQFAMELEQVTLTCEASGDPISITWRT-----STR 347
 173 QLELDSGTWTCTVLONOKKVEFKIDIVLAFQKASIIYKKEGVESSPPLATVEKL 232
 348 NISSEKKSMT---RPEKQETLDGHMVRSARVSSLTK---SIOT----- 388
 233 TGSSELIMQAEERASSSSKSWITFDLKNKEVSVKRVVODPRLQMGKPLPLHTLPALPOYA 292
 389 TDAGYVICTASNTIGQDS-----QSMYLEVGYPRLQO-----PVAV---YT 427
 233 GSGNLTALAEATGKGLHGVNLVWRATDLOKULICEVGGPSPKLMISLKLKNEAKYVS 352
 428 WEGN-----QVNITCEVFAYPS-----ATIS 448
 353 KKEKPVWLVNPEAGMWQCLSDSGOVLLESNIKVLPTWSTPVHPASALPAP 405
 449 -----WRFD-----GQLPSSNYSNIKITYT---PANSYLEVTP 479

RESULT 56
 T30581

neural cell adhesion molecule L1.1 - zebra fish (fragment)
 C:Species: Brachydanio rerio (zebra fish)
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
 C:Accession: T30581
 R:Tongioigi, E.; Bernhard, R.R.; Schachner, M.
 J. Neurosci. Res. 42, 547-561, 1995
 A:Title: Zebrafish neurons express two L1-related molecules during early axonogenesis.
 A:Reference number: 220875; MUID:96155762; PMID:8569941
 A:Accession: T30581
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1197 <TON>
 A:Cross-references: EMBL:X89204; NID:91065713; PID:91065714; PIDN:CAA61490.1
 C:Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology; i

Query Match 5.1%; Score 117.5; DB 2; Length 1197;
 Best Local Similarity 18.0%; Pred. No. 8.6;
 Matches 83; Conservative 84; Mismatches 122; Indels 123; Gaps 19;

9 HLLVLQALPLPAATGKNNKVLGKGGDTVELCTASQKKSIOFHHKNSNOIKLNGQSF 68
 259 HYTVVEAA--PYWTRSPREHLVAPGETVRLDCCADGIPAPNITW--SINGVPVSGTD--- 312
 69 LTKGPRSLNDRADRSRLMDGNFPLIKLIKIEDSDTYICEVDOKEEVQLVFGLTAN 128
 313 -----VDPRRRV---SSGKLITSNVFSDTAVYQCEAVNKGISILI-----N 351
 129 SDTHLQGSGLT---LESPPGSSPSVOCRS-----PRGKNIOGK 167
 352 THVHVEVPAQILTPDERLYQATAGQYVMDLCRTFGSPRLKIHWEILDSIPALNNAKISQ 411
 168 T-----LSVQLELDSGTWTCTVLONOKKVEFKIDV-----VLAFOKASSIVY 212
 412 TTNGSLKISNVSEDSNRKTVSVSTNKSISADVEVLAKRTYIGVPPONLHVIRSGDALIH 471
 213 KKEGEVFEFPLAFVTEKLTGSGELIMQAEERASSSSKSWITFDLKNKEVSVKRVYTOPDK 272
 472 CK-----YTVDHNLKSPYQVKNKGHITNS--TSNKKYHIEBSLKLVDVQM 517
 273 Q-MGKPLPLHTLPOLPQVAGSGLTLAEATGKL-----HGFVNLVWRATDLO 323
 518 EDMG-----IYSCVSTLDSPTAGYITVDKRPPOSILKI-----SEKME 559
 324 KNLICEVWGPRT---SPKLMISLKLKNEAKYVSREKPVWLVNPEAGMWQCLSDSGOVL 379
 560 RSVTIS--WPEVENSPTVEYIENBGE-----TDBEGMQKRVSYSDID 605
 380 LESNIKVLPTWSTPVHPASALPAPPTGSALPDPOTASALPD 421
 606 SMRSICYSKYNHQLQIRAVNSIGTAPTESSL-SYSTPAKRD 646

RESULT 57
 T43027
 neural cell adhesion molecule L1 - goldfish.
 N:Alternate names: E587 antigen
 C:Species: Carassius auratus (goldfish)
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
 C:Accession: T43027
 R:Giordano, S.; Laessle, U.; Lottspeich, F.; Stuenkel, C.A.O.
 submitted to the EMBL data library, April 1996
 A:Description: Molecular cloning of goldfish E587 antigen, a cell adhesion molecule expr
 A:Reference number: 222294
 A:Accession: T43027
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1232 <GIO>
 A:Cross-references: EMBL:U55211; NID:91305526; PID:91305527; PIDN:AAA9159.1
 C:Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology; i
 C:Keywords: cell adhesion; membrane protein

Query Match 5.1%; Score 117.5; DB 2; Length 1232;
 Best Local Similarity 21.5%; Pred. No. 8.9;

Matches 102; Conservative 68; Mismatches 173; Indels 131; Gaps 25;

Qy 2 NRGVFRLLLVLTQALPLPAATQGNKVVIGKKGDVLELTCTASQKSIQFMKNSNQKI 61
Db 286 SGGSVKHHYATVEAA--PYWTRRENILYAPGEIVRLDQAEIGPTNTITW--SMNGKPI 342
Qy 62 LGNGSFLTKGPPSKLNDADRSRSLMDQGNPPLIITKLIKEDSDTYICEVEDQKEEVL 121
Db 343 AGTD-----PDPFRHHV---SSGTLITLDVQI--SDPAVYHVEATMKGNI 383
Qy 122 VFGLTANSDTLH--LQGGSLT---LTLESPPGSSSVQCR---SPRCK----- 161
Db 384 I-----NTHVHVELEPPIITLEDLKYATEGQVTLQCRFGSPQAPVDMQITNSGPAL 438
Qy 162 -NIQGKT-----LSVSOLELQDSGTWCTVLQNGKVEFKIDIVLAFQKASSIYKKEG 216
Db 439 ANAKWSQTSDDNLOQSDVSEBSSMYTCSV--STSNMISALVYL--NRTKIYDPPQD 493
Qy 217 EGV---EFSFPLAFTVEKLTGSGELMWQAEARASSSKSWITFDLKNKEVSVKRVTDPKL 272
Db 494 LRVLRGDDAVLQCRVTVDHMLKQPTIQMKDKHKITTS-----ANDDKY 537
Qy 273 QMGKTLPLHLTPQALPQVAGSNTLTALAKTGKLGHEVNLVYWRATQLOKNTLCEVWG 332
Db 538 TESPDGSLKIT---DVQMEDSGIYSCISTKLDSVSAATGSIIV---LDK----- 580
Qy 333 PTSP-KLMLSLKLEN-----KEAKVSRERKRPVWVLPNPEAGMOCILSDS 375
Db 581 PGPSPHSLSEKKEKSVTLSMWPGAEENNSPISEVIERKEKQ---NGCKHMEYRRVP 636
Qy 376 GOVL-LESNIKVLPTWSTPVPVHPRASAL-----PAPRTGALPDPOGTASLPD 421
Db 637 QDITLHLIHLQPYSTY---HFRVARGVNGIGMSESPSPSESY---STPAKPD 682

RESULT 58

A32579
neuroglian - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: A32579
R/Biber, A.J.; Snow, P.M.; Horroch, M.; Patel, N.H.; Jacobs, J.R.; Traquina, Z.R.; Schi
Cell 59, 447-460, 1989
A>Title: Drosophila neuroglian: a member of the immunoglobulin superfamily with extensiv
A/Reference number: A32579; MUID:90030418; PMID:2805067
A/Accession: A32579
A>Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1239 <BIB>
A/Cross-references: GB:M28231; NID:9157998; PIDN:AAA28728.1; PID:9157999
A/Note: the authors translated the codon TAT for residue 1234 as Thr and AAA for residue
C/Genetics:
A/Gene: FlyBase:FBgn0002968
A/Cross-references: FlyBase:FBgn0002968
C/Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology; I
C/Keywords: alternative splicing; cell adhesion; duplication; membrane protein
F/353-412/Domain: immunoglobulin homology <IMM1>
F/446-502/Domain: immunoglobulin homology <IMM2>
F/535-596/Domain: immunoglobulin homology <IMM3>

Query Match 5.0%; Score 116.5; DB 1; Length 1239;
Best Local Similarity 19.5%; Pred. No. 10;
Matches 65; Conservative 59; Mismatches 131; Indels 79; Gaps 14;

Qy 25 GNKVVLGKGGDVTLELTCTASQKSIQFMKNSNQIKLIGNGSFLTKGPPSKLNDADRSR 84
Db 223 GNKVLLDSQNKPPVRYQVSRKQSLALGKMELEFCITYG---TPLPQTVWSR-DSQR 286
Qy 85 SLWD-----QGNF--PLIITKLIKEDSDTYICEVEDQKEEVLVFGLTANSDTLHQGS 138
Db 287 IOWSDRIQGHYKGLVIRQTNPFDAGTYCTDVSNGVNAQSFIILVNVSVPPFTKEPE 346
Qy 139 LTLTLESPPGSSPSVQCRS-----PRGNIQGK-----TLSVSOLELQ 177

Db 347 IATAED---EEVVFECRAQVPEPKISWTHNGPDIQSTPNPRRTVDNTIRINLVKG 403
Qy 178 DSGTWCTVLQNG---QKKEFKIDIVLAFQKASSIYKKEGQVFEFPLAFTVE-KL 232
Db 404 DTGNYGCAATNSAGYVYKDYLVNQAEPPITSEPAVSVTDGRNV-----TICRV 455
Qy 233 TSGSELMWQAEARASSSKSWIT---FDLK-NKEVSVKRVTDQPKLQMGKTLPLHLTPQAL 288
Db 456 NGSPKPLVKKLRAS---NMLTGRVYVQANGDLEIDVT----- 491
Qy 289 PQTAGSNTLTALAEAKTGKLGHEVNLVYWRATQI 322
Db 492 --FSDAGKYCYAQNKGEIGIADGSLVKEHTRI 523

RESULT 59

S49010
embryonic receptor kinase - mouse
C/Species: Mus musculus (house mouse)
C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 24-Sep-1999
C/Accession: S49010
R/Choi, K.; Wall, C.; Hanratty, R.; Keller, G.
Oncogene 9, 1261-1266, 1994
A>Title: Isolation of a gene encoding a novel receptor tyrosine kinase from differentiat
A/Reference number: S49010; MUID:94181281; PMID:8134130
A/Accession: S49010
A>Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1330 <CHO>
A/Cross-references: EMBL:X78568; NID:9510664; PIDN:CAAS5311.1; PID:9510665
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: Atp
F/823-1158/Domain: protein kinase homology <KIN>
F/831-839/Region: protein kinase Atp-binding motif

Query Match 5.0%; Score 116.5; DB 2; Length 1330;
Best Local Similarity 19.7%; Pred. No. 12;
Matches 91; Conservative 72; Mismatches 167; Indels 133; Gaps 21;

Qy 33 KGPVLELTCTASQKSIQFMKNSNQIKLIGNGSFLTKGPPSKLNDADRSRLMDQGNF 92
Db 568 EGEDLKSLCVN---KFLYRIDITWLTFTVNNRTMNSISKQNGHSSYSI---TL 617
Qy 93 PLIITKLIKEDSDTYICEV-----EPOKEEVLVFGLTANSDTLHQGSFLTLESPP 146
Db 618 NIVIKVNSLSDSGYACRANITYTGEDIKAKTEVLVRSRA---PHLLQNL---DYEVG 671
Qy 147 PGSSPSVQCRSPRG-----KN---IQ-----GKTLVSQLELQDSGTWCTCT 185
Db 672 ISGSTTLIDQQA-RGVPAPOITWLNKNHKLQOEBGIIIGPNSTLFIERVTEDEGVYRCR 730
Qy 186 VLQNGKVE-----FKIDIVLAFQKASSIY-----KKEGEVYER 221
Db 731 STNOKGAVESAAYLVQGISDKSNLELTITLYCVAAITLWLLTLFIRKLKSSSEVKT 790
Qy 222 SF-----PLAFTVEKLTGSGELM-WQAEARASSSKS-----ITFDLKN- 258
Db 791 DYLSIIMDPDEVLDEGCEHLPYDASWMEFARERLKLKSLGAGAFQKVQAQAFGIKGS 850
Qy 259 ---KEVSVKRVTDQPKLQMGKTLPLHLTPQALPQTAGSNTLTALAEAKTGKLGHEVNLV 315
Db 851 PTCRTVAVKMLKEGATSEYKALMTLEKILTHIGHILNVNVLGACTKQGGAPL-----NV 905
Qy 316 VMAATQVQKRLTCEWGPPTS PKLMLSLKENKKAQVSKREKRPVWVLPNPEAGMOCILSDS 375
Db 906 IVE-----YCK-IGNLSNYL-----KSRDLPCLNKDALKALHMKES 942
Qy 376 GOVLLESNIKVLPTWSTPVPVHPRASALPAP-TPGSAIPDPOGTAS 417
Db 943 LEPDLEGGCK-----PRLDVSSSSVTSSSPEDRSVS 975

```
RESULT 60
JC1509
biliary glycoprotein E - mouse
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 07-May-1999
C:Accession: JCI1509
R:McCuall, K.; Rosenberg, M.; Nedellec, P.; Turbide, C.; Beauchemin, N.
Gene 127, 173-183, 1993
A:Title: Expression of the Bgp gene and characterization of mouse colon biliary glycopro
A:Reference number: JCI1505; MUID:93273228; PMID:8500759
A:Accession: JCI1509
A:Molecule type: mRNA
A:Residues: 1-458 <MCC>
A:Cross-references: GB:X67280
C:Comment: This protein is expressed at the cell surface and plays a determinant role in
C:Genetics:
A:Gene: Bgpe
C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi
C:Keywords: glycoprotein; receptor
F:1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>
F:160-219/Domain: immunoglobulin homology <IMM1>
F:254-303/Domain: immunoglobulin homology <IMM2>
F:339-396/Domain: immunoglobulin homology <IMM3>
F:87,104,148,199,206,210,226,258,290,294,304,333,375/Binding site: carbohydrate (asn) (c
Query Match 5.0%; Score 116; DB 2; Length 458;
Best Local Similarity 19.6%; Pred. No. 3;
Matches 100; Conservative 76; Mismatches 185; Indels 148; Gaps 23;
QY 1 MNRG-VPRHLLLVQL-----ALLPAATQGNKVLGKKGDTVELTCTAS 44
DB 8 LHKGVPMVGLLTSLASLWSPPTTAETIAVPPQVADNNVLLVHNLPLALGA--- 64
QY 45 OKKSIQFHMKNSNQI-----KILGNQGSFLTGPSPKLANDRADRSLSMDQGNPLIIT 96
DB 65 -----FAMYKGNPVTNAEIVHVGVTGNTKT--TTGPAH-----SGRETVYNSGS--LLI 109
QY 97 KNLKIEDDTYICEVEDQ-----KEEVQ-----LLVFGLTANSPTHLQOGSLTTLTLES 145
DB 110 QKVYKDTGYVTIEMTDENFRRTAATVQPHVHPLLKPNITSNNSNPVEGDDSVSLTCD 169
QY 146 PPGSSPVQCRSPRGKNI-----OGSKTLSVSOLELQDSGTWCTCTVLQONQKVE-- 194
DB 170 YTDPNITVYLSMRNGBESLSEGRRLTSLTLLNTRNDTGPVCEP--RNPVSVNR 228
QY 195 --FKIDIV-----VLAFOKASSIYKKEGEQVESPPLAFTVEKLTSGGELMW---QAE 243
DB 229 DPFSLNIIYGPDPPII-----SPSDIYLHPGSNMLNLSCHAA-----SNPPAQYFWLINEX 279
QY 244 RASSKSWI-----TFDLKNKEVSVKRVTPQDPKLQMGKPLHLTLPLQALPYXA 292
DB 280 HASSQELFLPNITTTNNSGTYTCLVNNSVTGLSRTTVK----- 316
QY 293 GSGNLTALAEAKTGKLEHVNLVWRATOLQKRLTC-----EVMGPTSPKMLSLKLN 344
DB 317 ---NITV--LEPTVQPSLQVNTTVEKLDV--TLTCLSDNIGANIQWLFNSQSLQTERMTL 370
QY 345 ENKEAVSKREKPVWLVNPEAGMOCCLSDSGQVLLSEINIKVLPWTSTVPHRASALPAP 404
DB 371 TLSQNNSLIRIDP--IKREDAGEYQCEISNPVSVKRSNSIKL-----DIIFD 415
QY 405 PTGSALPDPTQTSALPDPPASALPALA 433
DB 416 PTGGSLSDGALAGIAGIVAGVALIAGLA 444
```

```
RESULT 61
S34338
biliary glycoprotein F - mouse
N:Alternate names: mouse hepatitis virus (MHV) receptor glycoprotein
C:Species: Mus musculus (house mouse)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
C:Accession: S34338; JCI1510; A41093
```

```
R:Huang, D.C.; Huang, X.F.; Novel, M.; Novel, G.
submitted to the EMBL Data Library, July 1992
A:Description: A Clp-family gene present on the lactose-protease plasmid of Lactococcus
A:Reference number: S34338
A:Accession: S34338
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-521 <HNU>
A:Cross-references: EMBL:X67281; NID:g312585; PIDN:CAA47698.1; PID:g312586
R:McCuall, K.; Rosenberg, M.; Nedellec, P.; Turbide, C.; Beauchemin, N.
Gene 127, 173-183, 1993
A:Title: Expression of the Bgp gene and characterization of mouse colon biliary glycopro
A:Reference number: JCI1505; MUID:93273228; PMID:8500759
A:Accession: JCI1510
A:Molecule type: mRNA
A:Residues: 1-61, 'Q', 83-141, 'P', 143-521 <MCC>
A:Cross-references: GB:X67281
R:Williams, R.K.; Jiang, G.S.; Holmes, K.V.
Proc. Natl. Acad. Sci. U.S.A. 88, 5533-5536, 1991
A:Title: Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen
A:Reference number: A41093; MUID:91288498; PMID:1648219
A:Accession: A41093
A:Status: preliminary
A:Molecule type: protein
A:Residues: 35-59 <MTL>
C:Comment: This protein is expressed at the cell surface and plays a determinant role in
C:Genetics:
A:Gene: Bgpf
C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi
C:Keywords: glycoprotein; receptor
F:1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>
F:160-219/Domain: immunoglobulin homology <IMM1>
F:254-303/Domain: immunoglobulin homology <IMM2>
F:339-396/Domain: immunoglobulin homology <IMM3>
F:87,104,148,199,206,210,226,258,290,294,304,333,375/Binding site: carbohydrate (asn) (c
Query Match 5.0%; Score 116; DB 2; Length 521;
Best Local Similarity 19.7%; Pred. No. 3.5;
Matches 100; Conservative 74; Mismatches 189; Indels 144; Gaps 22;
QY 1 MNRG-VPRHLLLVQL-----ALLPAATQGNKVLGKKGDTVELTCTAS 44
DB 8 LHKGVPMVGLLTSLASLWSPPTTAETIAVPPQVADNNVLLVHNLPLALGA--- 64
QY 45 OKKSIQFHMKNSNQIKILGNQGSFL-----TKRPSKLANDRADRSLSMDQGNPLIIT 98
DB 65 -----FAMYKGNPVTNAEIVHVGVTGNTKTTPAH-----SGRETVYNSGS--LLIQ 111
QY 99 LKIEDSDTYICEVEDQ-----KEEVQ-----LLVFGLTANSPTHLQOGSLTTLTLES 147
DB 112 VTVKDTGYVTIEMTDENFRRTAATVQPHVHPLLKPNITSNNSNPVEGDDSVSLTCD 171
QY 148 GSSPVQCRSPRGKNI-----OGSKTLSVSOLELQDSGTWCTCTVLQONQKVE-- 194
DB 172 DPNITVYLSMRNGBESLSEGRRLTSLTLLNTRNDTGPVCEP--RNPVSVNRSD 230
QY 195 FKIDIV-----VLAFOKASSIYKKEGEQVESPPLAFTVEKLTSGGELMW---QAE 245
DB 231 FSLNIIYGPDPPII-----SPSDIYLHPGSNMLNLSCHAA-----SNPPAQYFWLINEX 281
QY 246 SSSKSWI-----TFDLKNKEVSVKRVTPQDPKLQMGKPLHLTLPLQALPYXA 294
DB 282 SSQELFLPNITTTNNSGTYTCLVNNSVTGLSRTTVK----- 316
QY 295 GNLTLALAEAKTGKLEHVNLVWRATOLQKRLTC-----EVMGPTSPKMLSLKLN 346
DB 317 ---NITV--LEPTVQPSLQVNTTVEKLDV--TLTCLSDNIGANIQWLFNSQSLQTERMTL 372
QY 347 KEAVSKREKPVWLVNPEAGMOCCLSDSGQVLLSEINIKVLPWTSTVPHRASALPAP 406
DB 373 SQNNSLIRIDP--IKREDAGEYQCEISNPVSVKRSNSIKL-----DIIFD 417
QY 407 GSALPDPTQTSALPDPPASALPALA 433
```

Db 418 OGGSLDGAIAIGVIGVAGVALLIAGLA 444

RESULT 62

1JBONC

neural cell adhesion molecule short domain form precursor - bovine

N/Alternate names: NCAM-140

C/Species: Bos primigenius taurus (cattle)

C/Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text_change 22-Jun-1999

C/Accession: A32976; A38778; B44290; S05402

R/Lipkin, V.M.; Khrantsov, N.V.; Andreva, S.G.; Moshnyakov, M.V.; Petukhova, G.V.; Raki

FEBS Lett. 254, 69-73, 1989

A/Title: Calmodulin-independent bovine brain adenylate cyclase. Amino acid sequence and

A/Reference number: A32976; MUID:89378239; PMID:2776887

A/Accession: A32976

A/Molecule type: mRNA

A/Residues: 1-853 <LIP>

A/Cross-references: GB:X16451; NID:960; PID:CAA34470.1; PID:961

A/Accession: A38778

A/Molecule type: protein

A/Residues: 20-35;51-61;113-117;122-147;155-161;262-275;279-302;353-360;369-382;544-562;

A/Note: the authors identified this protein as calmodulin-independent adenylate cyclase

R/Rougon, G.; Marshak, D.R.

J. Biol. Chem. 261, 3396-3401, 1986

A/Title: Structural and immunological characterization of the amino-terminal domain of m

A/Reference number: A44290; MUID:86140120; PMID:3512556

A/Accession: B44290

A/Molecule type: protein

A/Residues: 20-36 <ROU>

A/Note: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol

C/Comment: Various forms of NCAM are produced by alternative splicing.

C/Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu

C/Keywords: alternative splicing; brain; cell adhesion; duplication; heparin binding; s

F/1-19/Domain: signal sequence #status predicted <SIG>

F/20-853/Product: neural cell adhesion molecule, short domain form #status experimental

F/20-719/Domain: extracellular #status predicted <EXT>

F/34-98/Domain: immunoglobulin homology <IMM1>

F/113-191/Domain: immunoglobulin homology <IMM2>

F/152-156/Region: heparin binding #status predicted

F/161-165/Region: heparin binding #status predicted

F/228-288/Domain: immunoglobulin homology <IMM3>

F/261-270/Region: NCAM binding #status predicted

F/321-396/Domain: immunoglobulin homology <IMM4>

F/428-480/Domain: immunoglobulin homology <IMM5>

F/527-604/Domain: fibronectin type III repeat homology <FN3A>

F/633-693/Domain: fibronectin type III repeat homology <FN3B>

F/720-737/Domain: transmembrane #status predicted <TM>

F/738-853/Domain: intracellular #status predicted <INT>

F/41-96;139-189;235-286;328-394;435-488/Disulfide bonds: #status predicted

F/222;314;346;432;458;487/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.0%; Score 116; DB 1; Length 853;

Best Local Similarity 20.5%; Pred. No. 6.9;

Matches 87; Conservative 52; Mismatches 159; Indels 126; Gaps 18;

32 KKGDVVELTCTASQKKSIOFMKNSNOIKIIGNGSFLTKGPKSLNDPDRBSRLMOGN 91

Db 130 REGDAVAVCDVSSLPPTIIMKHGRDVLTKDVRFTV-----LTN 171

Qy 92 PPLIINKLIEDSDTYICE-----VEDQKEVQLV-----FGLTANSDPHL 133

Db 172 NYLQIRGIRKKTDEGYRREGRIILARGEINFDIQIVVNPPTVGAROSIVATAN----- 226

Qy 134 LOGGSLTTLTSPSSPSVOCSPRGKNIQ-----GKTVLSQLELDQSGTW 182

Db 227 -LGSSVTLVCAAGPEPTVSW-TKDGEIENEDEKLFSDSSELTIRKDKNDKDEAY 284

Qy 183 TCTVLQNGKVEFKIDIVLAFQKASIVYKKEGQVFSFPLAFYVKLTGSGE----- 237

Db 285 VCIENKAGEODASIHLPKAPKPK--ITVVENQGTAMELEQVTLTCE--ASDPIPIST 338

Qy 238 LMMQERASSSK---SWTFDLKNKEVSVKRVTDPKLQMGKLLPL-HLTLPQALPOYAG 293

Db 339 TWRTSTRNISSEKASWTRE-----KQETLDGHMVVRSHARVSLTLKSI--QYTD 388

Qy 294 SGNLTALAEKTKGLHOEVNLVVMRATQL-----QKNLTCEVWGPTSPKMLLS 341

Db 389 AGEVYCTASNTIQDSQSMYLEVOYAPKLOGPAVYTWEGNQVITCEVAPPS----- 442

Qy 342 LKLENKAKYSKREKPVWLVNPEAGMMQGLSDSGOVLIESNKVLPTWSPVHPASAL 401

Db 443 -----ATIS-----WFRD-----GQLPSSNSNIKIYNT--PSASYL 473

Qy 402 PAPP 405

Db 474 EVTP 477

RESULT 63

139207

leukocyte surface protein V7 - human

C/Species: Homo sapiens (man)

C/Date: 02-Jul-1996 #sequence revision 02-Jul-1996 #text_change 21-Jan-2000

C/Accession: I39207

R/Ruegg, C.L.; Rivar, A.; Madani, N.D.; Zeitung, J.; Laus, R.; Engleman, E.G.

J. Immunol. 154, 4434-4443, 1995

A/Title: V7, a novel leukocyte surface protein that participates in T cell activation. I)

A/Reference number: I39207; MUID:95238941; PMID:7722300

A/Accession: I39207

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1021 <RES>

A/Cross-references: EMBL:233642; NID:9854194; PID:9854195

C/Genetics:

A/Gene: GDB:V7

A/Cross-references: GDB:702141

A/Map position: 1p13-1p13

C/Superfamily: human leukocyte surface protein V7

Query Match 5.0%; Score 116; DB 2; Length 1021;

Best Local Similarity 20.1%; Pred. No. 8.7; Indels 186; Gaps 27;

Matches 113; Conservative 69; Mismatches 195; Indels 186; Gaps 27;

Qy 16 LALLP--AATQGNKVVGLKKGDVVELTCTASQKKSIOFH-----W-----KNSNOIKIL 62

Db 140 LIVIPDTLSATWSQTLGKEGEPPLALTCASAKATQHTLSVTWYLTQGGSGQATEIL 199

Qy 63 GNGGSF-LTGPSPKANDRADSRSLMDQ--NPPLIINKLIEDSDTYICE---VD-- 113

Db 200 SLKDFLTVGPPYTERFAASDVQNLKGPTRFLSIERQSSDQGLFCFATEWIDPD 259

Qy 114 -----QKEVQLV-----FGLTANSDPHLQGGSLTTLT-LESPPGSSPSVQ--- 154

Db 260 ETWMTFKKQTDQTLRIQPAVDFOVNITRADSLPAGKLELVCLVSSGRDQLOGIW 319

Qy 155 -CRSPRGKNIQGGKTLV-----SOLEIQ-----DSGTWTC- 185

Db 320 FNGTEIAHIDAGVGLKNDYKERASQGLQSLKGPFAFLKIFSLGPEDEGAYRCV 379

Qy 186 -----VLQNGKVEFKIDI-----VTLAFQKASIVYKKEGQVEF----- 221

Db 380 AEVWKRTGSGWVYLRKQSPDSVHRLKPARSVVSTKKNQOVW--ESETLAFLCKAG 437

Qy 222 --SPPLAFT-----VEKLTGSGE-----LW-----MQERASSSKSWTFDLK 257

Db 438 GASSPLSVSMWHIIPRDTQPEFVAGMGQDGIVOLGALLMGTSYHGNTRELKMDATQLE 497

Qy 258 -----NKEVSVKRVTDPKLQMGKLL-----PLHLLPOLPOYASGN 296

Db 498 ITTFAITDSGYTCRVSSEKRNQARDLSWTQKISVTVKSLSSLSQVSLMSQPVMLTNT 557

Qy 297 LTTALAEKTKGLHOEVNLVVM-----RATOLQKLTCEVWGPTSPKMLLSKL 344

Db 558 FDLSCVVRAGYSDLKVPALTVTWQFOPASSHIFQLIRITHNGTIE--WGN-----FLSR 609

```

QY      345  ENKEAYKSRKREKVVWL-----NPEAGMOCILSDSGVLLESNTKYLPTMSTVPHRAS 359
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      610  FOKTKVKSGLPFSQQLVHDATEQELGVQLCEVEVYDRLNLYNN----- 653
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      400  ALPAPPTGSLPDP-QTASALPD 421
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      654  ---PPPASAIISHPLRIAVTLPE 673
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 64
S57845
protein-tyrosine kinase (EC 2.7.1.112), receptor type tie precursor - bovine
C/Species: Bos primigenius taurus (cattle)
C/Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 16-Jul-1999
C/Accession: S57845; S32690
R/Sato, T.N.; Qin, Y.; Kozak, C.A.; Audus, K.L.
Proc. Natl. Acad. Sci. U.S.A. 90, 9355-9358, 1993
A/Title: tie-1 and tie-2 define another class of putative receptor tyrosine kinase genes
/Reference number: S57845; MUID:94022374; PMID:8415706

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-1136 <SMT>
A:CrossRef-references: EMBL:X71423, NID:8296575, PIDD:CA50554.1, PIDD:8296576
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993
C:Superfamily: protein-tyrosine kinase, receptor type tie; EGF homology; fibronectin type
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-116/Product: protein-tyrosine kinase, receptor type tie #status predicted <MAT>
F:136-108/Domain: immunoglobulin homology <IM1>
F:213-253/Domain: EGF homology <EG1>
F:257-300/Domain: EGF homology <EG2>
F:304-342/Domain: EGF homology <EG3>
F:353-426/Domain: immunoglobulin homology <IM2>
F:447-528/Domain: fibronectin type III repeat homology <FN3A>
F:540-631/Domain: fibronectin type III repeat homology <FN3B>
F:640-728/Domain: fibronectin type III repeat homology <FN3C>
F:759-784/Domain: transmembrane #status predicted <TM>
F:835-1112/Domain: protein kinase homology <KIN>
F:843-851/Region: protein kinase ATP-binding motif
F:43-106,370-424/Disulfide bonds: #status predicted
F:84,159,501,554,707/Binding site: Asp hydrolytate (Asn) (covalent) #status predicted
F:868,885,977/Active site: Lys, Glu, Asp #status predicted

| | | | | | | | | | | |
|----|-------------|---|--------------|-------|------------|------|--------|--------|-------|----|
| | Query Match | Similarity | 5.0%; | Score | 115.5; | DB | 1, | Length | 1136; | |
| | Best Local | Similarity | 21.4%; | Pred. | No. 11; | | | | | |
| | Matches | 86; | Conservative | 47; | Mismatched | 146; | Indels | 123; | Gaps | 18 |
| Qy | 109 | CEVEDOKEEVOLLVGLTANSPDTHLLQGOSLTLTLESPPGSSP-----SVQCRSPRGKNI | 163 | | | | | | | |
| Db | 342 | CEKSRIPIQLDMVSELEFNLDT-----MPRINCAAGNPFPPVRSGMELRKXDDGYTL | 393 | | | | | | | |
| Qy | 164 | QQGKTL-----SVSQELQDSGTWTCTVLON--QKKVEFKIDIYLAFAQKASIV | 211 | | | | | | | |
| Db | 394 | LSTKAIVEPDRTAAFEVBRLADGGSLMECVSTSGGDSSRRFRINVPVPEPLTAARL | 453 | | | | | | | |
| Qy | 212 | YKKEEQVEFSPFLAFYVEKLTGSSELMMQARASSSKSWITFDLKNEVSVKRYTOPBK | 271 | | | | | | | |
| Db | 454 | LAKOSRQLVVSPVLVFSFGDPJIASVRLHRRP--DSTAMWST-----IIVDP- | 498 | | | | | | | |
| Qy | 272 | LQMKGKLPLHLTLPOALPOVAGSNVLTIA-LBAKTGLKHQENVLVVMRATOLOKULTECV | 330 | | | | | | | |
| Db | 499 | -----SEVNTLMNLRFKTKG---YSRVQLSRGEGEGE-----A | 529 | | | | | | | |
| Qy | 331 | WGPTS-----PKMLSLKENKEAKVSKREKPVVVLNDEAG-----MMQCLISD | 374 | | | | | | | |
| Db | 530 | WGPPLMTTTDCPEPLIKPWLEGMHVGEGRDLKVSMSLPVPQGPLVDGFILRLMD--GA | 586 | | | | | | | |
| Qy | 375 | SGQVLTLESNIK-----VLPTWST-----PYMRASALAPAPPGSAL | 410 | | | | | | | |
| Db | 587 | RGOERRE-NVSSPQARTALLTGLTGTGYTYOLDVRLHYCHTLGASPAPAAVL-LTPSGEPA | 644 | | | | | | | |

```

Oy      411 PDPQTASALPDP-----PASAALPALAVISFLLGIG 444
          ||| | | | | | | | | | | | | | | |
Db      645 PRLHAQALSDSEIQLMQRPAAAGPISKYIVEVQVAGSG 686

```

RESULT 65
A69392
hypothetical protein Afl138 - Archaeoglobus fulgidus
C|Species: Archaeoglobus fulgidus
C|Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C|Accession: A69392
R|Kleink, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
R|Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
G|Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
N|ature 390, 364-370, 1997
A|Authors: Uterbach, T.; Cotton, M.D.; Spriggs, T.; Arttach, P.; Kaime, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A|Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A|Reference number: A69250; MUID:98049343; PMID:9389475
A|Accession: A69392
A|Status: preliminary; nucleic acid sequence not shown; translation not shown
A|Molecule type: DNA
A|Residues: 1-554 <KLE>
A|Cross-references: GB:AEO01025; GB:AEO00782; NID:g2669348; PIDN:AAB90109.1; PID:g264945

[illegible]

N:Alternate names: connectin
 N:Contains: serine/threonine-specific protein kinase (EC 2.7.1.1.-)
 C:Species: Homo sapiens (man)
 C/Date: 12-Aug-1996 #sequence revision 12-Aug-1996 #text_change 15-Sep-2000
 C/Accession: 138344, 138345, S20897, S20899, S6365, S37393
 R/Labell, S.; Kolmerer, B.
 Science 270, 293-296, 1995
 A>Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.
 A/Reference number: A57430; MUID:96026330; PMID:7569978
 A/Accession: 138344
 A>Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
 A/Molecule type: mRNA
 A/Residues: 1-26926 <LAB>
 A/Cross-references: EMBL:X90568; NID:G1017424; PID:G1017425
 R/Musco, G.; Tiazzi, C.; Schuck, P.; Pastore, A.
 Biochemistry 34, 553-561, 1995
 A>Title: Dissecting titin into its structural motifs: identification of an alpha-helix
 A/Reference number: 138345; MUID:95119041; PMID:7819249
 A/Accession: 138345
 A>Status: nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1377-2014 <MUS>
 A/Cross-references: EMBL:X83270; NID:G602579; PIDN:CAA58243.1; PID:G602580
 A/Note: Conformation and properties are reported for a synthetic peptide corresponding to
 R/Labell, S.; Gautel, M.; Lahey, A.; Trinck, J.
 EMBO J. 11, 1711-1716, 1992
 A>Title: Towards a molecular understanding of titin.
 A/Reference number: S20897; MUID:92258380; PMID:1582406
 A/Accession: S20897
 A>Status: nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 13597-14200, 'I', 14202-14696 <LAB>
 A/Cross-references: EMBL:X64698; NID:G37192; PIDN:CAA5939.1; PID:G37193
 A/Accession: S20897
 A>Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: mRNA
 A/Residues: 16330-16362, 'S', 16384-16756, 'F', 16758-16860 <LAB>
 A/Cross-references: EMBL:X64699; NID:G37190; PIDN:CAA5940.1; PID:G37191
 A/Accession: S20899
 A>Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: mRNA
 A/Residues: 'P', 22278-22431, 'R', 22433-22448, 'G', 22450-22453, 'Q', 22455-22480, 'TR', 22483-2
 A/Cross-references: EMBL:X64697; NID:G37190; PIDN:CAA5938.1; PID:G37195
 R/Kolmerer, B.; Olivieri, N.; Witt, C.C.; Herrmann, B.G.; Labell, S.
 J. Mol. Biol. 256, 556-563, 1996
 A>Title: Genomic organization of M line titin and its tissue-specific expression in two
 A/Reference number: S6365; MUID:96177761; PMID:8604138
 A/Accession: S6365
 A>Status: nucleic acid sequence not shown
 A/Molecule type: DNA
 A/Residues: 26729-26825 <COL>
 A/Cross-references: EMBL:X92412; NID:G1236761
 R/Gautel, M.; Leonard, K.; Labell, S.
 EMBO J. 12, 3827-3834, 1993
 A>Title: Phosphorylation of KSP motifs in the C-terminal region of titin in differentiat
 A/Reference number: S37393; MUID:94008990; PMID:8404852
 A/Accession: S37393
 A/Molecule type: mRNA
 A/Residues: 26831-26926 <GAU>
 R/Improta, S.; Politou, A.S.; Pastore, A.
 submitted to the Brookhaven Protein Data Bank, February 1996
 A/Reference number: A66736; PDB:1TTT
 A/Contents: annotation; conformation by (1)H-NMR, residues 5253-5341
 R/Futtl, M.; Pastore, A.
 submitted to the Brookhaven Protein Data Bank, August 1996
 A/Reference number: A66201; PDB:1NCT
 A/Contents: annotation; conformation by (1)H-NMR, residues 'S', 26053-26155
 C/Genetics:
 A:Gene: GDB:TTN
 A/Cross-references: GDB:127867; OMIM:188840
 A:Map position: 2q31-2q32
 C/Function:
 A>Description: structural protein forming filaments in striated muscle

C/Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; prot
 C/Keywords: alternative splicing; calmodulin binding; cardiac muscle; duplication; glyco
 structural protein
 F:24752-25008/Domain: protein kinase homology <KIN>
 F:168,177,905,2276,2378,2451,2563,2659,2763,2896,3088,3179,3384,3432,3628,3772,4068,4
 98,11066,11488,11515,11635,11949,12170,12478,12526,12645,12875,13001,13036,13295,13540,1
 16780,16976,17579,17602,17667,17681,17845,17899,18121,18188,18209,18336,18670,18680,186
 21900,21935,22295,22495,22627,22897,23024,23318,23883,24012,24177,24290,24447,24642,248
 F:26171,26178,26184,26190/binding site: phosphate (Ser) (covalent) #status experimental
 Query Match 4.3%; Score 115; DB 1; Length 26926;
 Best Local Similarity 19.3%; Pred. No. 8.3e+02;
 Matches 115; Conservative 73; Mismatches 184; Indels 224; Gaps 26;
 QY 11 LVLQALPRA-TOG-NKVLGKGGDTVELTCTASQKKSIGFMKNSQIKLNGQS 67
 DB 14557 VIVEQTMLEPDLRGLGYQLVIAKAGDNKIKVPIVLGRKPTVTWKKQQLKQTRVN 14616
 QY 68 FLTKGPSKLNDRADSRSLWDQGNFLIKNLKTEDSDTYICEVED-----QKEEV 118
 DB 14617 FETTAISTININECVRS--DSGPPLTANIVGEVDVTTIQVHDIPGPTGPIKDEV 14674
 QY 119 QLLVGLTANSDFHLQGGSLVTLTLESPPGSSPSVQCRSPRGKNIQGGKTLG--VSQLEL 176
 DB 14675 -----SSD-----FVTFSDMPPEV-----DGVPISNVYVEMRQ 14703
 QY 177 QDSGTW---TCVLYQNKV-----EFKID-----IV----- 200
 DB 14704 TDSITWELATTVIRTTYKATRLTGLGVQFRVAKQNRGVGPGITSAMIVANYPKVP 14763
 QY 201 -----VLAFQKASSIVYKKEGVEFSPLATFVEKLTGSGELMWOAERA----- 245
 DB 14764 PRGTPVTAATKQSMITSMHEPLSDGSPRLTGHNVEKENGILMQVSKALVPGNIFKS 14823
 QY 246 SSSKSWITFDL-----KNKEVSVKRVTDPRYLQNGKKLPHL-----TLQALP 289
 DB 14824 SGLTDGIAVFRVIAENMAKSKRSPSEBMLADPDPGKRVPLNITHTVTLKWAAP 14883
 QY 290 QVAGSGNLTALAAK---TGKHHQ-----EVNLVYMAAQ-----LQKNLTCEV 330
 DB 14884 EYTGGEFKITIVYEKDDLPNGRWLKANFSNILEFTSGLTDAVAFVIAKNAAGAI 14943
 QY 331 MGPT-----SPKMLSLKLNK-----EAKVSKRPV--W----- 359
 DB 14944 SPSESEBDATTCDDVDPAKIKVDKFKVDVILKAGAFPLEADVSRPPTMEKSDGK 15003
 QY 360 -----VINPAGMWQCLSDSGQVLES-----NIKVLPTW 390
 DB 15004 ELEGTALEIKIADFTNLTNVKDS-----TRDSGAYTLTATNPGFAKHIIVNVKVLDR- 15057
 QY 391 STVHNPASALPAPPGSALPDPQTASALPDPRAASLPAALVIFSLDGLGIVA 446
 DB 15058 -----PQPPG-----PLAVTEVISEKCVLSWFPPLDQGA 15088
 RESULT 67
 S00682
 IGB:Fc receptor alpha chain precursor - human
 N:Alternate names: Fc-epsilon receptor
 C/Species: Homo sapiens (man)
 C/Date: 31-Dec-1988 #sequence revision 31-Dec-1988 #text_change 23-Jul-1999
 C/Accession: S00682; B30154; S42209
 R/Kochan, U.; Pettine, L.F.; Hakim, U.; Kishi, K.; Kinet, J.P.
 Nucleic Acids Res. 16, 3584, 1988
 A>Title: Isolation of the gene coding for the alpha subunit of the human high affinity Ig
 A/Reference number: S00682; MUID:88233953; PMID:2967464
 A/Accession: S00682
 A/Molecule type: mRNA
 A/Residues: 1-257 <KOC>
 A/Cross-references: EMBL:X06948; NID:G33317; PIDN:CAA30025.1; PID:G33318
 R/Shimizu, A.; Tepler, I.; Bentley, P.N.; Berenstein, E.H.; Stragatian, R.P.; Leder, P.
 Proc. Natl. Acad. Sci. U.S.A. 85, 1907-1911, 1988

A>Title: Human and rat mast cell high-affinity immunoglobulin E receptors: characterization
 A:Reference number: A94191; MUID:80150102; PMID:2964640
 A:Accession: B30154
 A:Molecule type: mRNA
 A:Residues: 1-257 <SH>
 A:Cross-references: GB:J03605; NID:9187449; PIDN:AAA36204.1; PID:9307164
 R:Yagi, S.; Yanagida, M.; Tanida, I.; Hasegawa, A.; Okumura, K.; Ra, C.
 Eur J. Biochem 220, 593-598, 1994
 A>Title: High-level expression of the truncated alpha chain of human high-affinity receptor product.
 A:Reference number: S42209; MUID:94170811; PMID:8125119
 A:Accession: S42209
 A:Molecule type: protein
 A:Residues: 26-197 <YAG>
 A:Experimental source: purified recombinant protein
 C:Genetics:
 A:Gene: GDB:FCER1A
 A:Cross-references: GDB:119902; OMIM:147140
 A:Map position: 1q23-1q23
 C:Superfamily: Fc gamma receptor III; immunoglobulin homology
 C:Keywords: immunoglobulin receptor; transmembrane protein
 F:1-25/Domain: signal sequence #status predicted <SIG>
 F:26-257/Product: 198 Fc receptor alpha chain #status predicted <MAT>
 F:44-95/Domain: immunoglobulin homology <IMM1>
 F:125-178/Domain: immunoglobulin homology <IMM2>

Query Match 4.9%; Score 114; DB 2; Length 257;
 Best Local Similarity 26.2%; Pred. No. 1.9;
 Matches 73; Conservative 34; Mismatches 92; Indels 80; Gaps 15;

```

QY 10 LLLVQLALLP-----AATGKNVVGK-----KGDVTELTCTAOKKSIOFHMKNSQI 59
DB 10 LLLVQLALLPAPGVLAVPQKPKVNSLPNNRPFKEBNTLTC-----51
QY 60 KILGSGSELTGPKSLNDRADSRSLDQGNFPLIKNLKIDSDTYICEV--DQKEE 117
DB 52 ----NGNNFEVSGTRKMFNG--SLSEETNSLNTVNAKEDSGEYKCHQGVNBSBP 103
QY 118 VQLVFGLTANSDFHLQGSITLTLESPPGSSPSVQCRSPGKRI-----QSGKITL-- 169
DB 104 VLEVF-----SDWLLQLQ-SAEVMEGQP--LFLRGCHGMNMDVYKVIYKQDEALKY 154
QY 170 -----SVSQLELDSGWTCT--VLONQKVEFKIDIVLVAFOKASSIYKKEGEV 220
DB 155 WENNNISTATVATVDSGYTGTGKVMQDYDESE-PLNTVTV--KA-----PREKYLQ 205
QY 221 FSPPLAFTVEKLTGSGELMWQAEARASSSKSWITFDLKNK 259
DB 206 FPIPLVILFAVDGLF-----ISTQQQVTFILKIK 237

```

RESULT 68
 C42632
 cell adhesion molecule apCAM (clone d12) - California sea hare
 C:Species: Aplysia californica (California sea hare)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 17-Mar-2000
 C:Accession: C42632
 R:Mayford, M.; Barzilai, A.; Keller, F.; Schacher, S.; Kandel, E.R.
 Science 256, 638-644, 1992
 A>Title: Modulation of an NCAM-related adhesion molecule with long-term synaptic plasticity.
 A:Reference number: A42632; MUID:92263095; PMID:1585176
 A:Accession: C42632
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-765 <MAX>
 A:Experimental source: CNS
 A>Note: sequence extracted from NCBI backbone (NCBIP:101351)
 C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immunoglobulin

Query Match 4.9%; Score 114; DB 2; Length 765;
 Best Local Similarity 18.1%; Pred. No. 8;
 Matches 65; Conservative 59; Mismatches 119; Indels 116; Gaps 17;

```

QY 29 VLAKKQDVELTCTASOKKSIOFHMKNSQIKILGNOGSFLTQKPSKLNDRADSRSLMD 88
DB 136 ILGSGE--EVECEVSGKPATVTWKEPNNKIDAGE-----KTTALN-----176
QY 89 QGNFPLIIKNLKIEDSDTYICEV-----EDQKEVOLLVFGL-----TANSDTHLLQ 135
DB 177 ----KLITKIDSLDSTKYLCDIIVDTGETKDFVIFTVVKLPITALPPIIHDPNKV- 231
QY 136 QGSILTLTLESPPGSSPSVQCRS---PRGNIOGKTLVSQLELDSGTTCTV-----186
DB 232 GDEKXITQQAQVGPDPPTQFQKGDVMTDEVVNNNG-VLTINPLKTTQATYTTCATYKKG 290
QY 187 -LONOKKVEFKI-----DIVLVAFOKAS--SIYKKEGEQVEFSPFL 225
DB 291 FAESSNLDVAVPPTIEMMETTYDAVSGELITICTAKGDEPESVINKKG-----PQ 343
QY 226 AFTVEKLTGSGELMWQAEARASSSKSWITFDLKNKESVYKRVTDPKLQMGKKLPLHLTP 285
DB 344 SASDGIYVKGPTV---EKVGSNNQ---DWEKTVV-----QHMTEFK 379
QY 286 QALQYVAGSGNLTALAEKTKLHGEVNLVYMRATOL-----QKNLTCE 329
DB 380 PV--TYQDAGTYICTAFSLVGSANKTVKLTQYKPNPDTDFKEREFGMRGKHALTQ 436

```

RESULT 69

C42632
 cell adhesion molecule apCAM (clone d15) - California sea hare
 C:Species: Aplysia californica (California sea hare)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 17-Mar-2000
 C:Accession: B42632
 R:Mayford, M.; Barzilai, A.; Keller, F.; Schacher, S.; Kandel, E.R.
 Science 256, 638-644, 1992
 A>Title: Modulation of an NCAM-related adhesion molecule with long-term synaptic plasticity.
 A:Reference number: A42632; MUID:92263095; PMID:1585176
 A:Accession: B42632
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-812 <MAX>
 A:Experimental source: CNS
 A>Note: sequence extracted from NCBI backbone (NCBIP:101346)
 C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immunoglobulin

Query Match 4.9%; Score 114; DB 2; Length 812;
 Best Local Similarity 18.1%; Pred. No. 8.7;
 Matches 65; Conservative 59; Mismatches 119; Indels 116; Gaps 17;

```

QY 29 VLAKKQDVELTCTASOKKSIOFHMKNSQIKILGNOGSFLTQKPSKLNDRADSRSLMD 88
DB 136 ILGSGE--EVECEVSGKPATVTWKEPNNKIDAGE-----KTTALN-----176
QY 89 QGNFPLIIKNLKIEDSDTYICEV-----EDQKEVOLLVFGL-----TANSDTHLLQ 135
DB 177 ----KLITKIDSLDSTKYLCDIIVDTGETKDFVIFTVVKLPITALPPIIHDPNKV- 231
QY 136 QGSILTLTLESPPGSSPSVQCRS---PRGNIOGKTLVSQLELDSGTTCTV-----186
DB 232 GDEKXITQQAQVGPDPPTQFQKGDVMTDEVVNNNG-VLTINPLKTTQATYTTCATYKKG 290
QY 187 -LONOKKVEFKI-----DIVLVAFOKAS--SIYKKEGEQVEFSPFL 225
DB 291 FAESSNLDVAVPPTIEMMETTYDAVSGELITICTAKGDEPESVINKKG-----PQ 343
QY 226 AFTVEKLTGSGELMWQAEARASSSKSWITFDLKNKESVYKRVTDPKLQMGKKLPLHLTP 285
DB 344 SASDGIYVKGPTV---EKVGSNNQ---DWEKTVV-----QHMTEFK 379
QY 286 QALQYVAGSGNLTALAEKTKLHGEVNLVYMRATOL-----QKNLTCE 329
DB 380 PV--TYQDAGTYICTAFSLVGSANKTVKLTQYKPNPDTDFKEREFGMRGKHALTQ 436

```

RESULT 70

A42632
cell adhesion molecule apCAM (clone d19) - California sea hare

C|Species: Aplysia californica (California sea hare)
C|Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 17-Mar-2000

C|Accession: A42632
R|Mayford, M., Barzilai, A., Keller, F., Schacher, S.; Kandel, E.R.
Science 256, 638-644, 1992

A|Title: Modulation of an NCAM-related adhesion molecule with long-term synaptic plasticity

A|Reference number: A42632; PMID:92263095; PMID:1585176

A|Accession: A42632
A|Status: preliminary; not compared with conceptual translation

A|Molecule type: nucleic acid
A|Residues: 1-932 <Max>

A|Experimental source: CNS
A|Note: Sequence extracted from NCBI backbone (NCBP:101342)

C|Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immunoglobulin-like domain

Query Match 4.9%; Score 114; DB 2; Length 932;
Best Local Similarity 10.1%; Pred. No. 10;
Matches 65; Conservative 59; Mismatches 119; Indels 116; Gaps 17;

Dy 29 VLKKKGDIIVELCTASQKRSIQFHWKSNOIKILGNGSPFTGPSKLNDRAISRSLMD 88
||:||::||::||:
Db 136 ILGGEG---EVECEVSGKPAPVTWKFENNKTIEGE-----XTIALN----- 176
||:||::||:
Dy 89 QGNPFLIIKNLKIEDSDTYICEV-----EDQKEVOLLVFGI-----TANSDTLLQ 135
||:||::||:
Db 177 ---KLTIKDLSLEDTKYKLCIDIVIDGETGFDFIDPTVVKLPTIALPPIIHDPKV- 231
||:||::||:
Dy 136 QSGLTLITLESPPGSSPSVOQRS----PRGNKIQQCKLTLSVGQLDLDSGMTTCIV----- 186
||:||::||:
Db 232 GDEVYITCGAVGPPEPYTFQFKGDVMTDEVMNNNG-VLTINPLKTTQATYTCLATNKGG 290
||:||::||:
Dy 187 -LONOKVTEFYI-----DIVLAFOQRAS--SIVVKSEGOVESFPFL 225
||:||::||:
Db 291 FAESNTLDVAVPPIETDEMERTVDASGGELLITICTAKGDEBPIWMKDG-----PQ 343
||:||::||:
Dy 226 AFVYEKLTVGSSELMMQAERASSSKSWITFDIKANREVSRYRTODPKLOMGKKLPRLHLLP 285
||:||::||:
Db 344 SASDTGIYNMKPGPTY---EKVSNQN---DMEEKIVA-----QHMTFFK 379
||:||::||:
Dy 286 QALPOLYASGNULTALEKTKLGHOENVNLVYMRRATL-----QKNIICE 329
||:||::||:
Db 380 PV--TYDDAGITYICTAFSLVGSANKTVKLTVQYXPNFDTDKEREFFGMGRGHKANLTQC 436
||:||::||:

RESULT 71

A30154
19E receptor alpha chain precursor - rat

N|Alternate names: FC-epsiloon-R alpha chain precursor
C|Species: Rattus norvegicus (Norway rat)
C|Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 21-Jan-2000

C|Accession: C31327; A31327; A30154; A27116; I55504
R|Lin, P.T., Albrandt, K., Robertson, M.W.
Proc. Natl. Acad. Sci. U.S.A. 85, 5639-5643, 1988

A|Title: cDNA heterogeneity suggests structural variants related to the high-affinity IgA binding site

A|Reference number: A94203; PMID:88289772; PMID:2965954

A|Accession: C31327
A|Molecule type: mRNA
A|Residues: 1-245 <Lit>
A|Cross-references: GB:M21622; GB:U03811
A|Experimental source: basophilic leukemia cell line, clone RJ-4

A|Accession: A31327
A|Molecule type: mRNA
A|Residues: 21-245 <Lit>
A|Cross-references: GB:M21622; NID:g204109; PID:NAAA41146.1; PID:g2041110; GB:U03811
A|Experimental source: basophilic leukemia cell line
R|Shimizu, A.; Teplov, I.; Benfey, P.N.; Berenshtien, E.H.; Straganiar, R.P.; Leder, P.
Proc. Natl. Acad. Sci. U.S.A. 85, 1907-1911, 1988

A|Title: Human and rat mast cell high-affinity immunoglobulin E receptors: characterization by complementary DNA sequencing

A|Reference number: A94191; PMID:86158102; PMID:2964640

A|Accession: A30154
A|Molecule type: mRNA

A:Residues: 1-245 <SH1>
A:Cross-references: GB:J03606; NID:g205331; PIDN:AAA1562.1; PID:g205332
R:Kinetic, J.P.; Metzger, H.; Hakiml, J.; Kochan, J.
Biochemistry 26, 4605-4610, 1987
A>Title: A cDNA presumably coding for the alpha subunit of the receptor with high affinity
A:Reference number: A27116; MUID:88024587; PMID:2959318
A:Accession: A27116
A:Molecule type: mRNA
A:Residues: 1,'G',3-236,'N',238-244,'RLKPNS' <KIN>
R:Repler, I.; Shimizu, A.; Leder, P.
J. Biol. Chem. 264, 5912-5915, 1989
A>Title: The gene for the rat mast cell high affinity IgB receptor alpha chain. Structure
A:Reference number: 155304; MUID:89174653; PMID:2522441
A:Accession: 155304
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-16 <RES>
A:Cross-references: GB:M25334; NID:g341335; PIDN:AAA4562.1; PID:g556391
C:Superfamily: Fc gamma receptor II; immunoglobulin homology
C:Keywords: immunoglobulin receptor; transmembrane protein
F:1-33/Domain: signal sequence #status predicted <SIG>
F:24-245/Product: IgE receptor alpha chain #status predicted <MAT>
F:142-93/Domain: immunoglobulin homology <IMW>

| | | | | |
|-----------------------|--------|---|----------------|---------------------|
| Query Match | 4.9%; | Score 113.5; | DB 2; | Length 245; |
| Best Local Similarity | 23.5%; | Pred. No. 1.9; | | |
| Matches | 63; | Conservative 41; | Mismatches 85; | Indels 79; Gaps 14; |
| OY | 11 | LVLVQLALLPATQGNKRVLTGK-----KGDVETLC--TASQKKSITQFHKWNSNOIKI | 61 | |
| Db | 13 | LVLISLGMLTATQKSVSLDPPIWIRILITGSKVLTICGNNSSQMNSYTK--WINDSI-- | 68 | |
| OY | 62 | LGNQGSFLTKGPKSLNBPADRSRSIMQGNPPLIIKNIKIDSPDYICEVED--OKEWQ | 119 | |
| Db | 69 | -----SNVKSSTM-----YVATITDSDGKYLQKGGFYKSKRY | 103 | |
| OY | 120 | LLVFGLTANSPTHLLOGOSLTLTLESPPGSSPVQCSPRG-----KNIQ | 164 | |
| Db | 104 | LNVM-----QEWLLQSSADVVLNDN--GSEFDIRCRWKKMKVHKVLYYKDIDIAFKSY | 154 | |
| OY | 165 | GGKTLVSQLELQDSGTWCTVLQONOKVEFKRIVLAFQKASIVYKKEGGEVSEFSP | 224 | |
| Db | 155 | DSNNISIRKAFPNDSGSHYCTGYLN--KVECKSDKFSIAYVKDYITERYR-----WQLIFP | 208 | |
| OY | 225 | ---LAFVTEKLTGSGELIMQAEPAASS | 248 | |
| Db | 209 | SLAVILFPAVD--TG---LWFSTHOFES | 231 | |

RESULT 72
 J05519
 50K glycoprotein precursor - chicken
 C|Species: Gallus gallus (chicken)
 C|Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 17-Nov-2000
 C|Accession: J05519
 R|Hancock, K.A.; Gooley, A.A.; Jeffrey, P.L.
 Mol. Brain Res. 44, 273-285, 1997
 A|Title: AvGP50, a predominantly axonally expressed glycoprotein, is a member of the IgLC
 A|Reference number: J05519; MUID:9725899; PMID:9073169
 A|Accession: J05519
 A|Molecule type: mRNA
 A|Residues: 1-338 <HAN>
 A|Experimental source: brain
 C|Comment: This protein belongs to the IgLCN's subfamily of cell adhesion molecules.
 C|Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-terminus
 C|Keywords: glycoprotein
 F.1-11/Domain: signal sequence #status predicted <SIG>
 F.32-338/Product: 50K glycoprotein #status predicted <MAT>
 F.46-113/Domain: immunoglobulin homology <IMM>
 F.146-199/Domain: immunoglobulin homology <IMM2>
 F.232-292/Domain: immunoglobulin homology <IMM3>
 F.40.136.148.279.287.300.315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.9%; Score 113.5; DB 2; Length 338;
Best Local Similarity 21.3%; Pred. No. 2.9; Indels 101; Gaps 20;
Matches 84; Conservative 63; Mismatches 146; Indels 101; Gaps 20;

10 LLLVQLALLPA-----TQGNKVVIGKKGVTELTCTASOKKSIQFMKNSNQIKI 61
14 LVLLRLCLLPGLPVRSDFTRGTDNITVAGDITAIRCFVEDRSS-KVAMLRSGIIF 72
62 LGNGSFLTKGPKSLNDPAD-SRSLMDQGNFPLIITKLIKEDSTTYICEVEDQ---KE 116
73 AG-----EDKMSIDPVELEKRSPLR---YSLRIQVDVDESSYCSYQTOQHHPKTS 122
117 EVOLLVPG-----LTANSPTHLLOGSLTLTLESPPGSSPSVQCR--SPRGNIOGKT-L 169
123 QYLVIVQVPKISNISITITVEGSSVTLVCANGRPREPVITWRHLITGKEPGESEYLL 182
170 SVSOLSDSGTWTCTVQONOKVFKIDIVLAFQKASSIYVKKEGQVE--PSFPLAF 227
183 EILGITRSGKYE-----KANEVASADVAVQVAVTVVYPTI 221
228 TVEK---LTSGGLMWOAERASSKSWITPDLKNKEVSVKRVITODPKLQNGKLPPLHLT 283
222 TESKSNBATGR---QALLPCEASAVPTPDF-----EWRDTRINSANGLEIKST 269
284 LPQALPOVAGSGNLTALAEAKTKHQEVNLVWMRATOLQKNLTCEVWGPTSPKMLM--- 340
270 -----GOSLWLVANVTEHYG-NITCVAMNLT-----GVTNASLYLYRP 308
341 -SLKLENKAKVSKREKPVVVLNPEAGMOCILS 373
309 GTGRVDNGSVSLA---VFLMLL---AASLLCLLS 336

RESULT 73
T42405
sax-3 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T42405
Rizalien, J.A.; Yi, B.A.; Bargmann, C.I.
Cell 92, 217-227, 1998
A:Title: The conserved immunoglobulin superfamily member SAX-3/Robo directs multiple asp
A:Reference number: Z22160; MUID:98117250; PMID:9458046
A:Accession: T42405
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1273 <ZAL>
A:Cross-references: EMBL:AF041053; NID:G2804779; PIDN:AAC38848.1; PID:G2804780
C:Genetics:
A>Note: SAX-3
C:Function:
A:Description: sax-3 function is required at the time of axon guidance

Query Match 4.9%; Score 113.5; DB 2; Length 1273;
Best Local Similarity 19.9%; Pred. No. 17; Indels 193; Gaps 24;
Matches 106; Conservative 59; Mismatches 175; Indels 193; Gaps 24;

2 NRGVFRHLVLVQLA-----LLPATQGNKVVIGKKGVTELTCTASOKKSIQF 51
3 NRKTLCLTLLVLAIVRSFCDASGNLAPVIEHIDIVVSRGSPATLNCGA-KPSTAKI 61
52 HMKNSNQIKILGNOSFLTGPSKLNDRADSRSLMDQGNFPLIIT---KLIKEDSTTY 107
62 TWYKDGQ--PVITNK-----EYVNSHRIYLDIGSLFLKLVNSGKNGKSDAGAY 108
108 ICEVEDQKEVOLVFGLTANSPTHLLOGSLTL-----TESPPGSSPSVQCR 155
109 YCVASNHEGVEK-----SNEG-----SLKLAMLREDPRVRRITVQALGGENAVLEC 154
156 RSPRG-----KNIOGKTLVSQLELQDSGTWTCTVQONOKV 193
155 SPFRGFPPEPVVSMRKDKELRIQDMPRYTLHSDGNLIIDPVDRSDSGTYQCV----- 206

194 EFKIDIVLAFQKASSIYVKKEGQVEFSPLAFTVEKLTGSGELMWOAERASSKSMIT 253
207 -----ANNV-----GERV--SNPARLSV-----FEKRFEBEPDMT 237
254 FDLKNKEVSVKRVITODPKLQ---GKKLPLHLTLPQALPOVAGSGNLTALAE----- 303
238 VDVGAVALPDCRVGDPOQPTWKRKNEPMPVT-----RAVIADKNRGLRIEVRQPSDEG 292
304 -----KTKLHQEVNLVWMRATOLQ-----KLTCEVWGPTSPK----- 337
293 EYVCYANPAGTLEASHLRVQAPSPQTRPADQSVAGTATFECTLVQQSPAYWMSK 352
338 -----LMLSLKENKAKVSKREKPVVVLNPE-----AGMOCILSDSGVLLBSNI 384
353 EGQODLLFPSTVYASDGTGTS---PYGTLIEVRQVDEGAYCAGMNSAGSSLSQAL 408
385 KVLPE---TWSTPVNPRASALPAPT-----GSALPPQTSALPDP 422
409 KVTTKAVTGNTPAKP-----PFTIEGHQNGTLMVGSSALIPQASGKPTP 454

RESULT 74
T30532
neural cell adhesion molecule L1 homolog - Fugu rubripes
C:Species: Fugu rubripes
C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 01-Mar-2002
C:Accession: T30532
R.Riboldi Tumnicliffe, G.R.; Platzer, M.; Nyakatura, G.; Elgar, G.S.; Brenner, S.; Rosen
submitted to the EMBL Data Library, September 1997
A:Description: Analysis of the genomic loci of Fugu rubripes homologs of the human disease
A:Reference number: Z20848
A:Accession: T30532
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1277 <RIB>
A:Cross-references: EMBL:AF026198; NID:G3098263; PID:G3098264; PIDN:AAC15880.1
C:Genetics:
A:Introns: 42/1; 47/1; 81/2; 149/1; 190/1; 247/1; 285/2; 347/1; 391/1; 440/1; 477/2; 531/
2
A>Note: L1-CAM
C:Superfamily: neural cell adhesion molecule L1, fibronectin type III repeat homology; in

Query Match 4.9%; Score 113.5; DB 2; Length 1277;
Best Local Similarity 21.7%; Pred. No. 17; Indels 81; Gaps 16;
Matches 80; Conservative 56; Mismatches 152; Indels 81; Gaps 16;

34 GDTVELTCTASOKKSIQFMKNSNQIKILGNOSFLTGPSKLNDRADSRSLMDQGNF 93
363 GETVKLDCQADGIRSPITW-----TVNGVPLSATSLSPRRSLTESGS-- 405
94 LIINKLIKEDSDTYICEVEDQKE-----VQLLVFGLTANSPTHLLOGSLTL-- 141
406 LIHKDVIFFGDTALIQCOASNKHGTTILANTVYVLELPPQILTEGNTYTFEGGKALLLEC 465
142 -TESP-----PSSPSVQCRSPRGNIOGKTLVSQLELQDSGTWTCTVQONOKVE 194
466 ETFGSPKPKTWESSISLTIADPRVNLITNG-GLEIANSYHDBEGIYITCLVQGSNISVN 524
195 FKIDI-----VLAFOKASSIYVKKEGQVEFSPLAFTVEKLTGSGELMWOAER---ASS 247
525 AEVEVNLNRVILNSPQALRL---QPKTALFTC-LVYTDPKLSPLQMKNDKXIFESH 580
248 SKSWITFDLKNKEVSVKRVITODPKLQNGKLPPLHLTLPQALPOVAGSGNLTALAEATGK 307
581 SDKKTYTFDGPGLIIS-----NVEPDGSGVYTCQIITIKLDWEVASTLTLC-DRDP 631
308 LHQEVNLVWMRATOLQKNLTCEVWG---TSPLKMLSLKENKAKVSKREK----- 356
632 VHLQVTVNAKRRVVTLN-----WTPGDNNNSPILYEVVEPDQMKENGWELKRVAAD 684
357 -----PVM 359
685 KKHVNLPLM 693


```
Db      216 RAIVDEFAEDNELT-----FKAGEIITVLDDSDPMMKGETHOGIGLFPSPNFV 264
Oy      229 VEKLTGSELWQARASSSKSMITFDLKNKEVSKRYTQDPK-----LQNGK 276
Db      265 TADLTA-----EPEMITKEKTVQF---SDVOVETIEPEEPAFIDEDKMDOLLQV-- 313
Oy      277 KLPLHLT-----LPQALPOYAGSGNLTLLAEKGTGLHQ-----EVNLVVMRATQL 322
Db      314 ---LQSTPSSDDQDPLPELHLHLMCHQMGPIDIEKLEDIDIKHSELSELVKNVMEALSL 370
Oy      323 QKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPEVWVLPNPEAGMOCILS--DSGOVLL 380
Db      371 YTKLMNE-----DPMYMWAKLQN-----QPYWQSGVSGSQYVAGPPSGAVLV 416
Oy      381 ESNITV--LPTWSTVHPRAS--ALPAPPTGSALPDQOTASALPD----- 421
Db      417 AGNAGMSHLQSYSLPEQLSLSLQAVVPSANPALPSOOTQAAVNTWVSSVOGNTYPSQ 476
Oy      422 -----PPASALPALAVISFLGLG 442
Db      477 APVSPPPAATAATAADVTLYQNG 502

RESULT 78
AC3616
succlnogllycan biosynthesis transport protein exop [imported] - Brucella melitensis (stra
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AC3616
R:DeJvecchio, V.G.; Kaparatl, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Leese
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AJ03252; PMID:11756688
A:Accession: AC3616
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-729 <KID>
A:Cross-references: GB:AE008918; PIDN:AAU54094.1; PID:q17965053; GSPDB:GN00191
A:Experimental source: strain 164
A:Gene: BMEI10852
A:Map position: 11

Query Match      4.9%; Score 113; DB 2; Length 729;
Best Local Similarity 20.9%; Pred. No. 8.8;
Matches 98; Conservative 65; Mismatches 155; Indels 150; Gaps 20;

Oy      82 SRSRLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVOLVGLTANSPTHLQGSQSLTL 141
Db      116 SRTVLSSQPDRTITLKNAR--DFDPAVRE---RQIDILLRLKVSND-----GRSFTI 164
Oy      142 TLESPPGSSPSVOQCSF-----RGKNIQGGKTLSSVSOLELQDSGTWTCTVLQNGKVEF 195
Db      165 -----FISYAPBPVYAAVANAFAATAYLV--HQVDYQGSAAARVSEMLGSKVLT 213
Oy      196 KIDI-----VLAFOKASSIVYKKEGEQVESFP---LAFTVEKLTSGGELMWQAEAS 246
Db      214 RNDLETAERVADEPFQKSLA---GEOGQISFOQRVWALTEIYVAATGAVSTQAARLQ 269
Oy      247 SSK-----SWITFDLKNKEVSVKRVNTQDPKLQWCK---LPLHLTLRQA 287
Db      270 TQAOLKNNNEAPAMTEILLASPAIONLKNDEARVQHLDLCKANGALKSAEIVLMAERES 329
Oy      288 LPQVAGSNLTLLAEKGTGLHQEVNLVVMRATQLOKNL----- 326
Db      330 LKQ-----QITAQVDEIITKSLSNEIRIAVQRTSLSEKELKEAETDLAKANGAQVRAQLD 384
Oy      327 -----TCVWGPTSPKMLSLKLENKEAKVSKREKPEVWVLPNPEAGM 367
Db      385 REANASRVVYETLYLTRYKQLEIQDGIAPLQALISQASPPAMAKASPR-----LIN----- 434
Oy      368 WQC-----LTSQGOVLLBSNIVKLPWTSPVHPRASALPAPPTGSALP--DPQ--TASA 418
```

```
Db      435 WLVLGLGIGLGFAPAGMTLRALDKIRPAQT-----ALALPGIPATLPLVPSQLTVPI 489
Oy      419 LPDPPASALP-----AALAVISFLGLGLGV 445
Db      490 LVNRGVDSASFPFGRAIKSVHDLRLALIRGRDLSALSVSSLEGGDKGL 537

RESULT 79
158164
BIG-1 protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 24-Sep-1999
C:Accession: 158164
R:Yoshinara, Y.; Kawasaki, M.; Tani, A.; Tamada, A.; Nagata, S.; Kagamiyama, H.; Mori, K
Neuron 13, 415-426, 1994
A:Title: BIG-1, a new TAG-1/F3-related member of the immunoglobulin superfamily with neu
A:Reference number: 158164; MUID:94338697; PMID:8060619
A:Accession: 158164
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1028 <RES>
A:Cross-references: EMBL:U11031; MID:9563132; PIDN:AAA63607.1; PID:9563133
A:Gene: Big-1
C:Superfamily: contactin; fibronectin type III repeat homology; immunoglobulin homology

Query Match      4.9%; Score 113; DB 2; Length 1028;
Best Local Similarity 19.8%; Pred. No. 14;
Matches 109; Conservative 69; Mismatches 174; Indels 198; Gaps 25;

Oy      23 TQGNKVVLGKK-----GDVYELTCTASQKK--SIQHHMKNKSNIKILNGSGSLTK--GP 73
Db      497 TEPRIILAPSNDVAVGESIILPCQVQHPDLDIMAFWFGTLTDFKXGSHFEKVG 556
Oy      74 SKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEV---QLTVFG----- 124
Db      557 SSSGD-----LMINIQKISGKVCVQVQVDSVSSAAELIVAGSPGP 601
Oy      125 -----LTANSPTH-----LQGSQSLTLT 142
Db      602 ENVKVDEITDTTQALSTEGTDSHSPISYAVQARTPSVGQWQVRTVPEAIDKTRTAT 661
Oy      143 -LESPP-----GSSPSVOQCSPRK-----NIOGKTLSSVQL--- 174
Db      662 VVELNPWEVEYEFVAVASNKIGGGEPSLPSEKVRTTEAAPVABSEVSGGG--SRSELVIT 720
Oy      175 -----ELQDSG-----TWCTVL---ONQKVEFKIDIVVLAFOKASSIV 211
Db      721 WDPVPELQNGGCGGYVAVRPLGVTTWICQVIVSPDNPRVFPNESIVPSPVEYKGV 780
Oy      212 YKKEGEVSESPFLAFTVEKLTSGGELMWQAEASSK--SWITFDLKN----- 258
Db      781 YNKGGEPPSPVTTVFSAEEPEPTVAPSHISHSLSSEIEVSWNTIIPWKSNGRLGVEV 840
Oy      259 -----KEYSVK--RTQDPKQLQMGKKLPHLTLPLQALPOY--AGSGNLTLLAEKGTG 307
Db      841 RYMNNGEESSESSVKAAGNQTSAVLRLGKSNLAVYTAVRVNTAGAGPSATVNAATTKK 900
Oy      308 LHQEVNLVVMRATQLOKNLTCEVWGPTSPKMLSLK---LENKEAKV-----SKRE 355
Db      901 -----TPSPQPPGNV---VNNATDTKVLNMBQVKALEN--ESEVTGKQVYKRTSSQ 947
Oy      356 KPVAVLPNPEAGMOCIL-----DSGOVLLBSNIVKLPWTSPVHPRASALP 402
Db      948 NNVOVLVNTNKTSAELLPIKEDIYIEVKATTDGDDGTSSEQIRI-----PRITSM 998
Oy      403 APPTGSALPD 412
Db      999 ARGSTASID 1008

RESULT 80
```



```

Oy 121 LVFGLTANSDPHLLOGGSLTLTLESPPGSSPSVQCRSPRG-----KN-----IOGKTL 169
Db 133 VVKVKQSSQSEPIHLQADPLETEKLMGECVVRSYEGNVTYKNGRVLQPVBEVVI 192
Oy 170 SVSOLELQDSG-----TWTCFTVL-----ONQKVEF----- 195
Db 193 NLARKVENSTGLFTWTSSLQWPTKEDANNAKFTCIIVTHGSPGQQTIOSEPVFVDVHYPT 252
Oy 196 -KIDVLAFOKASSIVVKKGEQVEFSPFLAFYTEKLTGSGE-----LWM---QARA 245
Db 253 EKVITRIVLS--QSSFTI-----KEGDNYTL-----KSGNGNPPQFLEFIPGETGI 298
Oy 246 SSSKSWITFDLKNKEVSVKRVTPDK-----LQMGKCLPLH 281
Db 299 RSSDPTVMTDVRNATGKCKSLINKSMMDATTIVHVLQDLTPSGEVTQOIGBALVVS 358
Oy 282 LTLQALP-----QVAGSGL-----TLAEAKTGKLTQEVNLVY- 316
Db 359 CTISSRNATVFWIKDNTRMKTSPSSSLQYODAGNYICETTLQVEGLKKRKTLLIVE 418
Oy 317 -----MRATOLQKLTCEVWGPTSPKMLSLKLENKAKVSKREKPVW 360
Db 419 GKPOIKMTKKTNTNKMSTIVCHVEG--FKPQAVQMTVYTGSGSLINKTEETKYV 470

```

RESULT 83

A58532

glial cell membrane glycoprotein LIG-1 precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Apr-1997 #sequence_revision 11-Apr-1997 #text_change 05-Nov-1999

C:Accession: A58532

R: Suzuki, Y.; Sato, N.; Tohyama, M.; Manaka, A.; Takagi, T.

J: Biol. Chem. 271, 22522-22527, 1996

A: Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically in

A: Reference number: A58532; MUID: 96394313; PMID: 8798419

A: Accession: A58532

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: mRNA

A: Residues: 1-1091 <SU2>

A: Cross-references: GB:D78572; NID:G1545806; PIDN:BA11416.1; PID:G1545807

C: Superfamily: leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan amino-ter

F: 36-61/Domain: proteoglycan amino-terminal homology <PMH>

F: 71-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F: 95-117/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F: 118-141/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F: 142-165/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F: 166-189/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F: 191-213/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F: 214-237/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F: 238-261/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>

F: 262-285/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>

F: 286-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>

F: 310-333/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>

F: 334-357/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>

F: 358-381/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>

F: 385-408/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>

F: 409-432/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>

F: 440-485/Domain: proteoglycan carboxyl-terminal homology <PCH>

Query Match 4.8%; Score 112; DB 2; Length 1091;
 Best Local Similarity 21.2%; Pred. No. 18;
 Matches 95; Conservative 59; Mismatches 166; Indels 120; Gaps 21;

```

Oy 14 LQALPLPAATGKKNVVGKGGDTVELTCTASQKSIQFHWKNSNOIKILNGSFLTYGP 73
Db 595 LTVNVLPSFTKIPHIAIRTTGTARLECAATGHPNPQIAMQDGG-----TDFP 643
Oy 74 SKLNRADRRSRSLMDQGNPPLIKNLKIEDSDTYICEVEDQKEVQLVFGLTANSDPHL 133
Db 644 A-----ARRRRNHWMPDDVFFITVDKIDMGVYCTAQNAGSV-----SANATLTV 691
Oy 134 LOGGSILTLTLES---PGSSPSVQCR-----SPRKNIGGKTLT----- 170

```

```

Db 692 LETPSLAFLPEDRVTVGEVAFQCKATGSPTRITWLKGGRLPLSTRHHFTPGNQLLV 751
Oy 171 VSOLELQDSGWTCTTVQONQKVEFKIDIVLAFQKASSIV-----YKGEQVEFSF 223
Db 752 VQNMIDAGGYTCETEMSN-----PLGTERHQSLSIILPTGCRDGTTVGI-F 798
Oy 224 PLATVEKLTGSGELMMQOERASSSSKSWITFDLKNKEVSVKRVTPDKLQMGKCLPLHLT 283
Db 799 TIAVCGSIVLTS--LWVVCIIYQTRK-----KSEESTYINDE-----TI 836
Oy 284 LPOLPOLYAGS-GNLTALAE--AKTGKLTQ-----EVLNVMRATYQLOKUL-----TCV 330
Db 837 VPPDPVSYLSQGTLSDRQETVVRTEGGHANGHIENGVCLRPDSLFPVEDIHSTTCR- 895
Oy 331 WGPSPPLMLSLKLENKAKVSKREKPVWVNLNPAAGMOCCLSPSGVLLESNTKVLPTW 390
Db 896 ----QPKLCVGYTRE--PWKTERADRTAAPHTTHAGSAVCSDC-----ST 936
Oy 391 STPVPRASALPAPPTGSALPDPTQASA 418
Db 937 DTAVHPC-----PVPDRSG---QPSTASS 957

```

RESULT 84

S46374

Ig kappa chain V-J region (T33-4) - human (fragment)

C: Species: Homo sapiens (man)

C: Date: 27-Jan-1995 #sequence_revision 01-Sep-1995 #text_change 21-Jan-2000

C: Accession: S46374; S38651

R: Bensimon, C.; Chastagner, P.; Zouali, M.

EMBO J. 13, 2951-2962, 1994

A: Title: Human lupus anti-DNA autoantibodies undergo essentially primary V(chi) gene re

A: Reference number: S46369; MUID: 94313975; PMID: 8039491

A: Accession: S46374

A: Molecule type: mRNA

A: Residues: 1-120 <BEN>

A: Cross-references: EMBL:Z27175; NID:G415965; PIDN:CAA81699.1; PID:G415966

C: Superfamily: immunoglobulin V region; immunoglobulin homology

C: Keywords: heterotetramer; immunoglobulin

F: 31-105/Domain: immunoglobulin homology <IMM>

Query Match 4.8%; Score 111.5; DB 2; Length 120;
 Best Local Similarity 34.6%; Pred. No. 0.99;
 Matches 37; Conservative 7; Mismatches 52; Indels 11; Gaps 3;

```

Oy 11 LTVQLALPLPA-----TQGNKVLGKKGDTVELTCTASQKSIQFHWKNSNOIKILGN 64
Db 2 LIGLILLMLPQAKCDVQGTQSPSTLSASVSGRVTITCRAGQSISTWLAWYQKS-----GN 57
Oy 65 QGSFLTGSPSKLNRADRRSRSLMDQG-NPPLIKNLKIEDSDTYICE 110
Db 58 APKLLIYGASVLESQVPSRFSGSGSDFTLTLSLQPDSDATYYCQ 104

```

RESULT 85

I37891

interleukin-11 receptor alpha chain - human

C: Species: Homo sapiens (man)

C: Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 03-Dec-1999

C: Accession: I37891; G01970

R: Chareil, M.; Sorel, M.; Leveau, B.; Dubois, S.; Moreau, J.F.; Bataille, R.; Minvielle,

Blood 86, 2534-2540, 1995

A: Title: Molecular cloning of two isoforms of a receptor for the human hematopoietic cy

A: Reference number: I37891; MUID: 95399754; PMID: 7670098

A: Accession: I37891

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: mRNA

A: Residues: 1-422 <RES>

A: Cross-references: EMBL:Z38102; NID:G995653; PIDN:CAA86224.1; PID:G995654

R: Van Leuven, F.

submitted to the EMBL Data Library, July 1995

A: Reference number: G08959

A: Accession: G01970

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-422 <VAN>
 A:Cross-references: EMBL:U32323; NID:g975334; PIDN:AAB36491.1; PID:g975335
 R:Van Leuven, F. EMBL Data Library, July 1995
 Submitted to the EMBL Data Library, July 1995
 A:Reference number: G08961
 A:Accession: G01971
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-422 <VA2>
 A:Cross-references: EMBL:U32324; NID:g975336; PIDN:AAB36492.1; PID:g975337
 C:Genetics:
 A:Introns: 34/1; 54/2; 111/1; 149/2; 160/2; 216/1; 270/3; 318/1; 358/1; 390/2; 418/1
 C:Superfamily: ciliary neurotrophic factor receptor; cytokine receptor homology; immunog
 F:120-310/Domain: cytokine receptor homology <CRS>

Query Match 4.8%; Score 111.5; DB 2; Length 422;
 Best Local Similarity 21.1%; Pred. No. 5.3;
 Matches 102; Conservative 53; Mismatches 174; Indels 155; Gaps 24;

QY 10 LLLVQLALL-----PATQGNKVVLLGKGDVETLC---TASQKSIQFHMKNNSQIK 60
 DB 11 VLVAVATLVASASSCPQAWGPPVQYQGPGRSVLCCPGVTAGDPVS---WFRDGEPK 66
 QY 61 ILNGGSFLTKGPSKLNDRADRSRLMDQG-NFPLIKNLKIEDSDTYICEVEDQKEVQ 119
 DB 67 LL-----QGP-----DSGLGHELVLAQADSTDEGTVCITQTLDG----- 99
 QY 120 LLVFGLTANSDTHLQSGSLTTLTLESPGSSPSVQCRPRGKNIQGGKTLVSQLELDS 179
 DB 100 -----ALGGTVTLQGLGPP-ARPVVSCQADYENF--SCTWSPSQI-----S 138
 QY 180 GTWCTVLONOKKVEFKIDIVLAFQKASIVYKKEGQVEFSPPLAFTVEKLTGSGELW 239
 DB 139 GLPRTYLYSYKKKTVLGAD-----SQRRS---ESTGWPPRQDPLG-AACVNHGAEPW 188
 QY 240 WQAEKASSSKSWITFDLKNKEVSKRVQ-DPKLQMGKPLHLTLPGALPOVAGSGNLT 298
 DB 189 SQYRINTEVNPPLGASTRLDVSLSILRPP-----PQGL----- 224
 QY 299 LALEKTKGLQEVNLYVMRATQLOKMLTCEVWGPTSKMLSLKLENKAKVSKREKV 358
 DB 225 -----RVESVPGYPRRLRASWTYPASWCCQPHFLKRLQYRPA-----QHFA 267
 QY 359 W-VLNP-----EAGMQCLSDSQVLLAESNI-KVLPWTM 390
 DB 268 MSTVPRAGLEVITDAVAGLPHAVASARDPLDAGTGTWSPPEANGTPTSTGTIKKELPAM 327
 QY 391 -----STVPHRASALPAPPTGSALPDPQTASALPDPAPASALPAPALAVISFL-----L 439
 DB 328 GOLHTQPEVEPQVDS-PAPPRPSLQPHRLLDH-RDSVEQVAVALASIGLIFLGLVAGAL 385
 QY 440 GLGL 443
 DB 386 ALGL 389

RESULT 86
 S41051
 fibroblast growth factor receptor-2 - eastern newt
 C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
 C:Date: 19-Mar-1997 #sequence_revision: 19-Mar-1997 #text_change 21-Jan-2000
 C:Accession: S41051
 R:Poulin, M.L.; Chiu, I.M.
 Biochim. Biophys. Acta 1220, 209-211, 1994
 A:Title: Nucleotide sequences of two newt (Notophthalmus viridescens) fibroblast growth
 A:Reference number: S41050; MUID:94146117; PMID:8312364
 A:Accession: S41051
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-750 <POU>
 C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote

F:101-162/Domain: immunoglobulin homology <IM>
 F:408-693/Domain: protein kinase homology <KIN>

Query Match 4.8%; Score 111.5; DB 2; Length 750;
 Best Local Similarity 21.6%; Pred. No. 11;
 Matches 79; Conservative 41; Mismatches 102; Indels 143; Gaps 16;

QY 15 QLLALPPATQGNKVVLLGKGDVETLCCTASQKSIQFHMKNNSQIK 72
 DB 94 KLVAVPAA-----NTVKPRCPAGCNTPPSRMVLKQKEFKQEHRIIGFKYRSQ 141
 QY 73 -----PSKLNDRADRSRLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVOLLV 122
 DB 142 HPSLIMESVVP-----DEGNATCIMEENYGSINHTYHLDVVERSPHRDILQ 188
 QY 123 FGLTANSDT-----HL-----LQGSLLTTLTLESPGSSPSVQCRS 157
 DB 189 AGLPANNTTKVGGDAEFCVKVSDAOPHIQWIRHPELNGSKI-----GPDGHPYLVK 242
 QY 158 PRGKNIQGGKTLVSQLELDQSGTWC-----TVL-QNOKKVEFKIDIV 201
 DB 243 RSGINSSNAEVLTLHNTEADAGQYTCVSNYIGEANQSAWLVLPASEKDERELD--- 299
 QY 202 LAFQKASSIY-----KKEGEQVEFSPPLAFTVEKLTGSG 236
 DB 300 -----SSSEYETAIYCVGFLITGMITGIMVCHMKGRKSDFFSPPA--VHKLSKSL 351
 QY 237 ELWMQ-----AERASSKS-----WITFDLKNKE-----VSVKRVYQDP-----KIQ 273
 DB 352 PLRQVTVASDSSSSMNSNTPLVYRITRLLSSNNDTHLAGVSEYELDEPKMEYPREKLT 411
 QY 274 MGKKL 278
 DB 412 LKPL 416

RESULT 87
 S03517
 T-cell receptor gamma chain precursor (clone pm17664) - human
 C:Species: Homo sapiens (man)
 C:Date: 07-Sep-1990 #sequence_revision: 07-Sep-1990 #text_change 21-Jan-2000
 C:Accession: S03517
 R:Tighe, L.; Forester, A.; Clark, D.M.; Boylston, A.W.; Javentir, I.; Rabbitts, T.H.
 Eur. J. Immunol. 17, 1729-1736, 1987
 A:Title: Unusual forms of T cell gamma mRNA in a human T cell leukemia cell line: implic
 A:Reference number: S03517; MUID:88083067; PMID:2961573
 A:Accession: S03517
 A:Molecule type: mRNA
 A:Residues: 1-340 <TTG>
 A:Cross-references: EMBL:X06774; NID:g37346; PIDN:CAA29941.1; PID:g37347
 A>Note: this sequence was determined from the differentiated gene
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: T-cell receptor
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-340/Product: T-cell receptor gamma chain #status predicted <MAT>
 F:21-116/Domain: V region (V-gamma-8) #status predicted <VR8>
 F:121-135/Domain: J region (J-gamma-2) #status predicted <JRB>
 F:136-340/Domain: C region (C-gamma-2) #status predicted <CRS>
 F:160-225/Domain: immunoglobulin homology <IM>

Query Match 4.8%; Score 111; DB 2; Length 340;
 Best Local Similarity 20.9%; Pred. No. 4.3;
 Matches 77; Conservative 58; Mismatches 138; Indels 96; Gaps 16;

QY 10 LLLVQLALLPATQGN-----KVVLGKGDVETLCCTASQKSIQFHMKNNSQIK 62
 DB 3 LALALLAFLPPASQKSSMLEGRTKSVTRPTGSAVITCDLPENNVYTHM-----YL 55
 QY 63 GNQGSFLTKGPSKLNDRADRSRLMDQG-----NFPLIKNLKIEDSDT 106
 DB 56 HQEG---KAPQRLLYDSDNSRVL-ESGISSEKHTYASTGSKLFIENLIERSGV 110
 QY 107 YICEVEDQKEVOLLVFGLTANSDTHLQSGSLTLT---LESPPGSSPSVQCRSPRGKNI 163

```

Db      111 YCATWTDRYYKKL-FG-----SGTLVTVTDKQLDADVSPKPTIFLP----- 152
Qy      164 OGCKLUSVQLELDQSGTMTCTVLQNOKKVEFKIDIVLAFQ--KAASIVYKKEGE----- 217
Db      153 -----SIAETKLOKAGTYLCIL-----EKFFPIIKIKHMOEKSNITLGGQEBNTTKT 200
Qy      218 ---QVEFSPLAFYTEKLTGSGELMMQAEARASSKSMITFDKKNKEVSKVATODPKLQW 274
Db      201 NDTYKFSW-LTVPEBSLDKEHC--IYRHENKNGIDQEIIFPIKTDVTTVPKQSY 256
Qy      275 GK-----KLPRLTLPOALPOYAGSNTLALAEATGKLHOEVNLVWKA 319
Db      257 SKDANDVTVDPKYVNSKDANDVTMTDPKDNWSKDANDTLILLQNTSAVYVYLLLLKS 316
Qy      320 TOLQKLTG 328
Db      317 VVYFAITTC 325

```

RESULT 88

cell adhesion molecule gp80 - slime mold (Dictyostelium discoideum)
 A44100
 C.Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
 C.Accession: A44100
 R.Desbarats, L.; Lam, T.Y.; Wong, L.M.; Siu, C.H.
 J. Biol. Chem. 267, 19655-19664, 1992
 A.Title: Identification of a unique CAMP-response element in the gene encoding the cell
 A.Reference number: A44100; MUID:92406924; PMID:1326559
 A.Contents: AX2
 A.Accession: A44100
 A.Status: preliminary
 A.Molecule type: nucleic acid
 A.Residues: 1-514 <DES>
 A.Note: sequence inconsistent with the nucleotide translation
 A.Note: sequence extracted from NCBI backbone (NCBI:P.113993)

Query Match 4.8%; Score 111; DB 2; Length 514;

Best Local Similarity 22.0%; Pred. No. 7.5; Matches 96; Conservative 60; Mismatches 181; Indels 100; Gaps 22;

```

Qy      63 GNQGSFLTGPSPKLNDRADRSRLMDQGNFPLIKNLKIEDS-DYICEVEDEQKEVQL 121
Db      116 GRIGMLINDGPNISG-----YKLNVDNSINAMLSV--TADSVSPT 155
Qy      122 VEGLTANSPTLLQGSLLTLESPPGSSPSVQCRS-----PRGNIQ 164
Db      156 IYPLVNT-----IAGGLNLLELIOFGFSTIVTSKVSFSPITTSITPLAFDLTPNNTVT 211
Qy      165 GG---KILSVSQ-----LELDGRTWCTVLQNOKKV-EFKIDIVLAFQKASSIYK 213
Db      212 GRYFVTTASVTMGSHIYTGLTIVQDDGT-NCHVIFTRSYESSNITAKASTGVDMYILD 270
Qy      214 KEQGEVFEFSF---PLAFTVEKLTGSGELMMQAEARASSKSMITFDKKNKEVSKVATQD 269
Db      271 NQGNQOPITFTYNPPTITSKQVNDSEI-----STNTGTDG---TQSLTMGTSS 319
Qy      270 PK-LQMGKRLPLHLLPOLAPQVAGSNTLALAEATGKLHOE---VNLVWKAATOLQK 324
Db      320 PNLVITGTNEKIVITLPHALPE---GRIGNLKAGISNVVTSITLVTPVINSTVQAPH 375
Qy      325 N-LTEWVGPTSPKMLSLKENKEAKV-----SKREKVVNLAEAGCMQCLSDSQV 378
Db      376 NQGSITISGIFLNNAHVSIVDQNTTDICADSGESITICVDAGSGITNLVLTNKNF 435
Qy      379 LLESNIKVLPTWSTPV-----HPRASALPAP-PTGSALPDP-----OTASALPDPAPAS 426
Db      436 ASDPITKTATSTYTTITDTPPTDTPATPSPTPTDTPATPSPTPTSTPEETAPSSAT 495
Qy      427 AL-PAALAV--ISPLL 439
Db      496 TLISPLSLIVIFISVL 512

```

RESULT 89

A45254
 surface glycoprotein BEN precursor - chicken
 C.Species: Gallus gallus (chicken)
 C.Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 21-Jul-2000
 C.Accession: A45254; S19202
 R.Pourquie, O.; Corbel, C.; Le Caer, J.P.; Rosier, J.; Le Douarin, N.M.
 Proc. Natl. Acad. Sci. U.S.A. 89, 5261-5265, 1992
 A.Title: BEN, a surface glycoprotein of the immunoglobulin superfamily, is expressed in
 A.Reference number: A45254; MUID:92302224; PMID:1608932
 A.Accession: A45254
 A.Status: preliminary
 A.Molecule type: mRNA
 A.Residues: 1-588 <POU>
 A.Cross-references: EMBL:X64301; NID:963087; PIDN:CAA45579.1; PID:963088
 C.Keywords: glycoprotein

Query Match 4.8%; Score 111; DB 2; Length 588;

Best Local Similarity 19.2%; Pred. No. 8.9; Matches 91; Conservative 68; Mismatches 175; Indels 140; Gaps 19;

```

Qy      8 RHLLVLQIALLPATQGNKRVLGKGPVVELTCTASQKSIQF-HMK---NSNQIKIL 62
Db      16 RPLCLLALACMPALGLYTVNAVYGDITMPCRLVDPDLMGKMKYEMPNSPVPIA 75
Qy      63 GNQGSFLTGPSPKLNDRADRSRLMDQGNFPLIKNLKIEDSDYICEV--EDQKEVQL 120
Db      76 FRSS---TKKNVQYDVPDYDKDRLSLSENYTLISIKNARISDEKRFVCMLVTEDDVSEPT 132
Qy      121 LVFGLTANSPTLLQGSLLTLESPPGSSPSVQCRSPG---KN-----IQGKTL 169
Db      133 VVKVFKQSPQPEILHQADPLETEKLMGECVVDSDSPBGNVTYKGRVLQVBEVVYI 192
Qy      170 SVSLELDQSG-----TWCTVL---QNOKVEF----- 195
Db      193 NLRKVENRSTGLFTMTSSLQYMPFKEDANAKFTICIVYTHGSGQKTIQSEVVDVNHPT 252
Qy      196 -KIDIVLAFQKASSIYKKEGEVFEPLAFTVEKLTGSGE-----LW---QAEKA 245
Db      253 EKVTIRVLS--QSGTI---KEGDVNTL---KCSGNGNPPQEFFIYIPGETEGI 298
Qy      246 SSKSMITFDKKNKEVSKVATQDPK-----LQMGKRLPLH 281
Db      299 RSSDVTYVTDVRRNATGEYKCSLIDKSMDDTTTTHVLDQLTPSGEVTYQIGBALPVS 358
Qy      282 LTLPOALP-----QYAGSNTL---TLAEATGKLHOEVNLV- 316
Db      359 CTISSNRATVFWIKDNTRMKTSFSSLOQYQDAGNYICETTHGEVGLKRRKTKLILVE 418
Qy      317 -----MRATOLQKNTLCEVWGPTSPKMLSLKENKEAVSKREKPVWV 360
Db      419 GKPOIKMTKNTNMKSTIVCHVEG--FPKPAQVMTVTSGLSILNTEETKYV 470

```

RESULT 90

C81412
 NOL1/NO2/sun family protein Cj0636 [imported] - Campylobacter jejuni (strain NCTC 11168)
 C.Species: Campylobacter jejuni
 C.Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 C.Accession: C81412
 R.Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Baham, D.; Chillin
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Batrel
 Nature 403, 665-668, 2000
 A.Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
 A.Reference number: A81250; MUID:20150912; PMID:1068824
 A.Accession: C81412
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-279 <PAR>
 A.Cross-references: GB:AL139075; GB:AL111168; NID:96967817; PIDN:CAB75272.1; PID:9696810
 A.Experimental source: serotype O2, strain NCTC 11168


```

0Y 23 :QGGKVVVLGKK-----GDTVELTCTASQKK--SIQFMKXSNIKILGNGSLTK-GR 73
Db 497 TEPRRIILAPSNMDVAGESVILPCVOVHDPDLIDMFMYFNGLTDFKDGSHFEKYG 556
0Y 74 SKLNDRADRSRLMDQGNFPLIIRKLKIEDSDTYICEVEDOKEV-----OLTVG----- 124
Db 557 SSSGD-----LMIRNIQLKHSKKYCWVQGTVDVSSAAELIVGSGRP 601
0Y 125 -----LTANSDFH-----LLOGQSILTT 142
Db 602 ENVKVDEITDTTAOLSTWGTGTDHSPVISAVOARTPFSVGMSGRVTPEVIDGKHTAT 661
0Y 143 -LESPP-----GSSPSVOCRSRPRCK-----NIOGKTLSTSLQ--- 174
Db 662 VELNPWVEYFPRIVASNKIGGSEBSLPSEKRVETEEAAPETAPESVSGGG-SRSEILVTT 720
0Y 175 -----ELODSG-----TWTCTVL---ONOKVYEFKIDIVLAFKASSIV 211
Db 721 MDVPPEELQNGGGGCGYVAFRRPLGVTWITQIVTVSPDNPRVFRPNESIVPSPRYEVKGV 780
0Y 212 YKKEGEQVESFPPLAFIVETKLTGSGELMWOAERASSK--SWITFDLK----- 257
Db 781 YNNKGEGCFPSVTTVFSAESEPTVAPSHISHSLSSEIEVSNMTIIPKLSNGHLGVEY 840
0Y 258 -----KNEVSXKV-----TODPKLQMGKTLPLHLTPOLAPQY--AGSGULTLALAKT 305
Db 841 RYWNNGGEEESRKYKAGNGTSAVLRG--AKSNLAAYTAARAANSAGAGFSAFSAVATTT 898
0Y 306 GKLHQEVLVVMRATQLOKNTLCEVWGPTSPKLMSS--LKLVEKKEAKV-----SKR 354
Db 899 KK-----TPSQQPPGNV---VMNATIDPKVLLINEQVAMENESSVTCYKFFYTTSS 946
0Y 355 EKPVWVLNPEAGMOQULL-----SDSGOVLLESNIIKVLPTWSTVPRASAL 401
Db 947 QNNVHVLTNKTSAELLPLPIKEDIYIEVKATITDGGDGTSSQIRI-----PRITSM 997
0Y 402 PAPRPGSAL 410
Db 998 DARGSTSAI 1006

```

A:Result 94
 A:11228
 Protein-tyrosine kinase (EC 2.7.1.112) Flk-1 precursor; endothelial cell-specific receptor
 C:Species: Mus musculus (house mouse)
 C:Date: 19-Jun-1992 #sequence revision 19-Jun-1992 #text_change 04-Feb-2000
 C:Accession: A41228; A46065; I56365; S18832; S29391
 R:Matthews, W.; Jordan, C.T.; Gavin, M.; Jenkins, N.A.; Copeland, N.G.; Lemischke, I.R.
 Proc. Natl. Acad. Sci. U.S.A. 89, 9026-9030, 1991
 A:Title: A receptor tyrosine kinase cDNA isolated from a population of enriched primitive
 A:Reference number: A41228; MUID:92020984; PMID:1717995
 A:Accession: A41228
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1367 >NMT>
 A:Cross-references: GB:X59397; NID:G50976; PIDN:CA42040.1; PID:G50977
 R:Millauer, B.; Witzigmann-Voos, S.; Schunuch, H.; Martinez, R.; Moller, N.P.; Risau, W.
 Cell 72, 835-846, 1993
 A:Title: High affinity VEGF binding and developmental expression suggest Flk-1 as a major
 A:Reference number: A46065; MUID:93208880; PMID:7681362
 A:Accession: A46065
 A:Status: Preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-24, 'T', 26-782, 'VL', 785-916, 'C', 918-1367 <ML>
 A:Cross-references: GB:X70842; NID:G57923; PIDN:CA550192.1; PID:G57924
 A:Note: Submitted to the EMBL Data Library, January 1993
 A:Note: Sequence extracted from NCBI backbone (NCBI:128064)
 R:Oelrichs, R.B.; Reid, H.H.; Bernard, O.; Ziemiecki, A.; Wilks, A.F.
 Oncogene 8, 11-18, 1993
 A:Title: NYK/Flk-1, a putative receptor protein tyrosine kinase isolated from E10 embryonic
 A:Reference number: I58365; MUID:93141255; PMID:8423988
 A:Accession: I58365

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-678, 'D', 680-1340, 'RSPV', <OE>
A:Cross-references: GB:S53103; NID:g264004; PIND:AA25043.1; PID:g264005
C:Genetics:
A:Gene: FLK-1; NYK
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
F:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; transmembrane
F:830-1165/Domain: protein kinase homology <KIN>
F:838-846/Region: protein kinase ATP-binding motif

Query Match 4.7%; Score 109; DB 2; Length 1367;
Best Local Similarity 21.0%; Pred. No. 38;
Matches 45; Conservative 36; Mismatches 73; Indels 60; Gaps 9

```

QY 9 HLLVLQGLALLPAAQGNKKVVLGKKGDVYELCTSAQKKSIOFHWNQNSQIKILNQGSF 68
Db 544 HVIRKPEITVPAQPTQO-----EVSLLCTADRNTFENILTWYK-----LGSQATS 590
QY 69 LTKGSPKINDRADSRSLW-----DOGNPPLII--KNLKIEDSDTVICEVEDQKEE- 117
Db 591 VHMBES-LTPCKKNDALMKLNGTWFNSSTNDIILVAFQNALQDQGDVCSAQDKKTKK 649
QY 118 -----VOLLVFGELTANSDDTHLLQGSLLTLLTESPGSSPSVQCRSPBGXN----- 162
Db 650 RHCLVKGQILIERMAPMITGNLENGTTI-----GETIEVTC--PASGNPPTHITWFXD 701
QY 163 -----IQGKTLVSGQLELDQSGTWC 184
Db 702 NETLVEDSGIVLRDGNRNILTRRYKEDQGGILVTC 735

```

RESULT 95
150478
neurolin - goldfish (fragment)
C/Species: Carassius auratus (goldfish)
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996
C/Accession: I50478
R/Laessing, U.; Giordano, S.; Stecher, B.; Lottspeich, F.; Stuermer, C.A.
Differentiation 56, 21-29, 1994
A/Title: Molecular characterization of fish neurolin: a growth-associated cell surface protein
A/Reference number: I50478, MUID:94299040, PMID:8026643
A/Accession: I50478
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-523 <LAb>
A/Cross-references: GB:L25056, NID:g407318, PID:g407319

| | | | | |
|-----------------------|-------|--|--|--------------------------------------|
| Query Match | 4.6% | Score 108 | DB 2 | Length 523 |
| Best Local Similarity | 10.7% | Pred. No. 12 | | |
| Matches | 73 | Conservative | 68 | Mismatches 127; Indels 122; Gaps 20. |
| QY | 34 | GDTEVLTCTAOKKS--- | IOFHMKNSNOIKLTGNQGSFLTGKPSK----- | LNMRADSR 84 |
| DB | 2 | GETTVLPENDGCTKKP | DGLIFTRMK--YVKDGGSPGDLIVQAQDEATVSATDGYKSRV | 58 |
| QY | 85 | SLMOQGNPLIIIKULKIEDSDTYICEVEDQEEVQLVFGLTANSDFTLGOSLTFILE | | 144 |
| DB | 59 | SI--AANSLLIAAGSLADQRFVTC----- | MYVSFTN-----LEYSVEGVKH | 99 |
| QY | 145 | SPGSPSPVQCRSPRGKNIQCGKTLVSQLELODSGTWTCTVLONQKVEFKDIIVLAF | | 204 |
| DB | 100 | KKP--SAPYIK--NNAKELENGKLTQJGECVEANA----- | | 130 |
| QY | 205 | QKASSIYKKEGEBOV---- | EFSPLPFAVTEKLTGSGELMWQAEPASSKSMITFDLKNK | 259 |
| DB | 131 | NPPADLLMKNNQNLVLDGKTIITTSITTKOKING----- | LSSTSRLOGTATAKE | 180 |
| QY | 260 | EVSRYKRYQDPKLTQMGKTLPHLTLPLQALPQYAGSGNLTLLAEKTKGLH-- | QEVNLVY | 316 |
| DB | 181 | DVE--SQFCTAKHWG--PDQVSEPSFP----- | IHPTEKVSILQV | 218 |

Qy 317 MRATOLQK-----NLTCVWGPSTPKMLSLKLENKAVSKREKPVWL-----NPEAGMW 368
 Db 219 VSQSPTRRGEVDYTLKCGQADGNPP---TSFNNFKIGKKVYTKDQVYTLTGVNADSGIT 275
 Qy 369 QCLSDSGQVLES-----NIKVLPT 389
 Db 276 KCSLLDND--VWESQFVTSPLDVSILPT 303

RESULT 96

168093
 C:Species: Homo sapiens (man)
 C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jan-2000
 C:Accession: 168093
 R:Berle, F.; Dubreuil, P.; Mattei, M.G.; Devillard, E.; Lopez, M.
 Gene 159, 267-272, 1995
 A:Title: The human PRK2 gene, related to the human poliovirus receptor gene (PVR), is th
 A:Reference number: 153960; MUID:95347610; PMID:7622062
 A:Accession: 168093
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-538 <RES>
 A:Cross-references: GB:S79172; NID:91042204; PID:91042205
 A:Gene: PRK2delta
 C:Superfamily: poliovirus receptor; immunoglobulin homology
 F:276-331/Domain: immunoglobulin homology <IMM>

Query Match 4.6%; Score 108; DB 2; Length 538;
 Best Local Similarity 21.3%; Pred. No. 13;
 Matches 112; Conservative 58; Mismatches 204; Indels 152; Gaps 25;
 Qy 2 NRGVFRLLLVQLALPAATQGNV-----VIGKKGDTVELTC----- 41
 Db 10 SRSPETPLMLPLLLLLLETTGAQDVAVLPEVRQGLGTVELPCHLLPVPGLYISLVT 69
 Qy 42 -----TASOKKSIQFMKNSNQI---KILNQSGFLTKGSKLNDRASSRLMDQNF 92
 Db 70 WQRPAPANHQNVAHFHKMGSPSPKPSERLSFVS---AKOSTGDTAEALDA--- 123
 Qy 93 PLIINKLIEDSDTYICEVD--QKEEVO---LIVFGILTANSDFHLLQGSLTTLTLESP 147
 Db 124 TLALHGLTVEDEGNTCEFAFPFKGVSFGMTLRIYAKPKN-----QAEAKVTFSQDP 177
 Qy 148 GSSPEVQGRSPRGK---NIQSGKTLVSQLELDOSGTCVYLQNKVKEKIDIVLAF 204
 Db 178 --TVALCTISKGRPPARISWLSLIDWEAKETQVSGTLAGTV---TVTSRFTLVPSGR 230
 Qy 205 QKASSIVYKKEGEQVE--FSFPLAFVE-----KLTGSGELMWQAE-----RASSS 248
 Db 231 ADGVTVTKVHESEFEPEPALIPVLTLSVRYPREVSISSGIDDMWYLGRTATATLSCVRSNP 290
 Qy 249 KSWITFDLKNKEVSVKRYTQDPKLOMGKPLH-----LTLPOAL----- 288
 Db 291 PTGYDMSTTSGTFPSAVALQ-----GSLVTHAADSLENTTFVCTVNAVAGMGRAEQVI 344
 Qy 289 -----FOYAGSGNLTLEAKTKGKHQEVNLVWKAATOLQKNLTC----- 328
 Db 345 FVRETPNTAGAG-----ATGGIIGIITAIATVAATGILLCRQGRKEQTLQGAED 397
 Qy 329 -EVMGPTSPKMLSLKLENKAVSKREKPVNLINPEAGMQCLLS---DSGQVLTLEENI 384
 Db 398 EDLEBPP-----SYKPPTPKAKLEAQEMPSQLFTLGSEHSPLKTPYFDAGASTCEQEM 451
 Qy 385 ---KVLPTW---STVHPAPASALPAPPTGSALPDFQTASALPDP 424
 Db 452 PRYHELPTLEERSGPIHGAATSL-----GSPIPV-----RGP 486

RESULT 97
 S40370
 Ig kappa chain - human

C:Species: Homo sapiens (man)
 C:Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
 C:Accession: S40370
 R:Klein, R.; Jaenichen, R.; Zachau, H.G.
 Eur. J. Immunol. 23, 3248-3271, 1993
 A:Title: Expressed human immunoglobulin chi genes and their hypermutation.
 A:Reference number: S40312; MUID:94080891; PMID:8258341
 A:Accession: S40370
 A:Status: preliminary; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-122 <KLE>
 A:Cross-references: EMBL:X72480; NID:9441428; PID:CAAS1148.1; PID:9441429
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotrimer; immunoglobulin
 F:30-104/Domain: immunoglobulin homology <IMM>

Query Match 4.6%; Score 107.5; DB 2; Length 122;
 Best Local Similarity 35.5%; Pred. No. 1.6;
 Matches 38; Conservative 5; Mismatches 53; Indels 11; Gaps 3;

Qy 11 LVIVQLALPA-----TQGNKVLGKKGDTVELCTASOKKSIOFHKNSNQIKILGN 64
 Db 1 LIGLLMLKAGACDILQITQSPSSLASVGDRAVYTCRASQISTFLHWQON---LGK 56
 Qy 65 QGSFLTKGPKSLNDRADSRSLMDQG--NFPILIKNLIEDSDTYICE 110
 Db 57 APKLIRYAANLQSGVPSRPSGSGSDFTLTISGLQPEDFATYYCQ 103

RESULT 98

S68177
 C-CAM2a protein isoform precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
 C:Accession: S68177
 R:Lucka, I.; Cichocka, I.; Baemler, K.; Bechler, K.; Reutter, W.
 Eur. J. Biochem. 234, 527-535, 1995
 A:Title: A short isoform of carcinoembryonic antigen-related rat liver cell-cell adhesion
 A:Reference number: S68177; MUID:96128184; PMID:8536699
 A:Accession: S68177
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-458 <LUC>
 A:Cross-references: EMBL:X91137; NID:91160272; PID:CAA62577.1; PID:91160273
 F:1-138/Domain: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-terminal
 F:1-33/Domain: signal sequence #status predicted <SIG>
 F:34-458/Product: C-CAM2a protein isoform #status predicted <MAT>
 F:252-301/Domain: immunoglobulin homology <IMM>

Query Match 4.6%; Score 107.5; DB 2; Length 458;
 Best Local Similarity 19.2%; Pred. No. 11;
 Matches 101; Conservative 80; Mismatches 175; Indels 169; Gaps 28;

Qy 3 RG-VFPRHLLLVQL-----ALPAATQGNKVLGKKGDTVELCTASOK 46
 Db 10 RQGIPLWRGLLTLTSLTYWSPLTQAQTVDAVPPNVVEKSVLL-----LAHNPQE 61
 Qy 47 KSTOFHKNSNQIKILANQSGFLTK-----GPKSLNDRADSRSLMDQNPPLIKLKI 101
 Db 62 FOV-FYVYKGTLL-----NPDSEIARYIRSDNMSKTGPVAGRETIYSNGS--LFFQVNVK 114
 Qy 102 EDSPTVYCEVEDQ-----KEVQLVVE-----GLTANSPTHLLOGGSLTTLTLESPGSS 150
 Db 115 TDERAYTLVDFDOQFNFIQTSVQFRVYALQKNVNGN--NSNPWEGEPF--VSLMCEPYTN 172
 Qy 151 PSYQCRSPRKNL-----OGKTLVSQLELDOSGTCVTLQON---QKVEERKID 198
 Db 173 NTGYLWNRNBSLSEGRVYFSSGNRTLTLNVRRTDKGYEECARPATFNRRSDPNLD 232
 Qy 199 IV-----VLAFOKASSIVYKKEGEQVEFSF---PLA---FTYEKL--TGSGELMWQA 242
 Db 233 VIVGPDAVVI-----SPDILVHQSGNINLSCHADSNPAPQYFWLINELKQTSQSELFTSN 288

| | | | |
|----|-----|--|-----|
| Qy | 243 | BRASSSKSMWTFEDLKN-----KEVSRYKRTQPKQMGKKLPLHLTLPPALPOYA | 292 |
| Db | 289 | ITTNNSGYVACE-VNNVTYGLSRTTVKNIYFEPVTPQSITQITNTVKEI----- | 337 |
| Qy | 293 | GSNGI/LTALAEK-TGKTLHQEVNLVVMRATQLOKNI/LTCEVMGPSTPKML-----SLKLENK | 347 |
| Db | 338 | --GSVTLTLCFCKDYG-----VSVR-----WLFNSQSIQ/LDRMTLSQDNS | 375 |
| Qy | 348 | EAKYS--KREKPVWVLNPEAGMMQCLSDSCQVLLSESNIKVLPWSTFPVHPRASALPAP | 405 |
| Db | 376 | TLRIDPIKRE-----DAGDYCEISN----- | 397 |
| Qy | 406 | TGSA/LPDPQTASALPDPPPASALPALAAVISFLLGLGAGVACVLA | 450 |
| Db | 398 | VSFRIHPDKIDVYIPDPPTQNSGISEGAIAGIVSAGVAAFLIA | 442 |

RESULT 99

A44783

ecto-ATPase precursor - rat

N:Alternate names: canalicular bile acid transport protein

C:Species: *Rattus norvegicus* (Norway rat)

C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 23-Jul-1999

C:Accession: A44783; A44410

R:Lin, S.H.; Guidotti, G.

J. Biol. Chem. 264, 14408-14414, 1989

A:Title: Cloning and expression of a cDNA coding for a rat liver plasma membrane ecto-ATPase

A:Reference number: A44783; MUID:89340561; PMID:2527235

A:Accession: A44783

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-519 <LIN>

A:Cross-references: GB:J04963; NID:G203989; PID:AAA41104.1; PID:G203990

R:Stipe, C.J.; Suchy, F.J.; Ananthanarayanan, M.; Perimutter, D.H.

J. Biol. Chem. 268, 2083-2091, 1993

A:Title: The rat liver ecto-ATPase is also a canalicular bile acid transport protein.

A:Reference number: A44410; MUID:93131966; PMID:8420979

A:Accession: A44410

A:Molecule type: protein

A:Residues: 110-120;122-138;148-150 <STP>

C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-terminal

C:Keywords: glycoprotein; liver; phosphoprotein; transmembrane protein

F:1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>

F:160-217/Domain: immunoglobulin homology <IMM2>

F:252-301/Domain: immunoglobulin homology <IMM3>

F:337-394/Domain: immunoglobulin homology <IMM3>

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:50:44 ; Search time 6.2213 Seconds
(without alignments)
3791.557 Million cell updates/sec

Title: SEQ5
Perfect score: 2325
Sequence: 1 MNRGVRPHLLVLVLQALLP.....VISFLGLGAGVACIARTR 453

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 125 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------|
| 1 | 2032 | 87.4 | 458 | 1 | CD4_HUMAN |
| 2 | 2008 | 86.4 | 458 | 1 | CD4_PANTR |
| 3 | 1858 | 79.9 | 458 | 1 | CD4_MACFU |
| 4 | 1857 | 79.9 | 458 | 1 | CD4_MACMU |
| 5 | 1854 | 79.7 | 458 | 1 | CD4_MACFA |
| 6 | 1849 | 79.5 | 458 | 1 | CD4_MACNE |
| 7 | 1832 | 78.8 | 458 | 1 | CD4_CERAE |
| 8 | 1728 | 74.3 | 397 | 1 | CD4_ERYPA |
| 9 | 1718 | 73.9 | 397 | 1 | CD4_CERTO |
| 10 | 1582.5 | 68.1 | 457 | 1 | CD4_SAIISC |
| 11 | 1147 | 49.3 | 459 | 1 | CD4_RABIT |
| 12 | 1137 | 48.9 | 463 | 1 | CD4_CANFA |
| 13 | 999 | 43.0 | 457 | 1 | CD4_RAT |
| 14 | 993 | 42.7 | 457 | 1 | CD4_MOUSE |
| 15 | 293 | 12.6 | 240 | 1 | CD7_HUMAN |
| 16 | 162 | 7.0 | 739 | 1 | VCAL_HUMAN |
| 17 | 155 | 6.7 | 739 | 1 | VCAL_HUMAN |
| 18 | 145 | 6.2 | 761 | 1 | NCA2_HUMAN |
| 19 | 144 | 6.2 | 739 | 1 | VCAL_MOUSE |
| 20 | 141.5 | 6.1 | 1493 | 1 | NEOI_MOUSE |
| 21 | 141 | 6.1 | 6885 | 1 | SNE2_HUMAN |
| 22 | 139 | 6.0 | 1447 | 1 | DC2_MOUSE |
| 23 | 138 | 5.9 | 837 | 1 | NCA2_MOUSE |
| 24 | 137.5 | 5.9 | 702 | 1 | CEA5_HUMAN |
| 25 | 135 | 5.8 | 1259 | 1 | SHS1_RAT |
| 26 | 134.5 | 5.8 | 509 | 1 | SHS1_RAT |
| 27 | 134.5 | 5.8 | 1260 | 1 | CAML_MOUSE |
| 28 | 131 | 5.6 | 1257 | 1 | CAML_MOUSE |
| 29 | 130 | 5.6 | 1906 | 1 | KMUS_CHICK |
| 30 | 127 | 5.5 | 3707 | 1 | PGM2_MOUSE |
| 31 | 125.5 | 5.4 | 725 | 1 | NCA2_MOUSE |
| 32 | 125.5 | 5.4 | 1447 | 1 | DCC_HUMAN |
| 33 | 125.5 | 5.4 | 4391 | 1 | PGM1_HUMAN |

| | | | | | | |
|-----|-------|-----|------|---|-------------|---------------------|
| 34 | 125 | 5.4 | 349 | 1 | LACH_SCHAM | Q26474 echinocerc |
| 35 | 125 | 5.4 | 2012 | 1 | DSCA_HUMAN | O60469 homo sapien |
| 36 | 124 | 5.3 | 1115 | 1 | NCA1_MOUSE | P13596 mus musculu |
| 37 | 124 | 5.3 | 1298 | 1 | YTFN_HAEIN | Q57523 haemophilus |
| 38 | 123 | 5.3 | 837 | 1 | NCM2_HUMAN | O15394 homo sapien |
| 39 | 122.5 | 5.3 | 564 | 1 | C166_BRARE | Q90460 brachydantio |
| 40 | 122.5 | 5.3 | 848 | 1 | NCA1_HUMAN | P13591 homo sapien |
| 41 | 121.5 | 5.2 | 333 | 1 | AMAL_DROME | P15364 drosophilila |
| 42 | 121.5 | 5.2 | 338 | 1 | LAMP_HUMAN | Q13448 homo sapien |
| 43 | 121 | 5.2 | 348 | 1 | KILO_RAT | Q92018 ratu |
| 44 | 121 | 5.2 | 997 | 1 | SCPI_RAT | Q03410 ratu |
| 45 | 120.5 | 5.2 | 773 | 1 | PIGR_RABIT | P01832 oryctolagus |
| 46 | 120.5 | 5.2 | 1091 | 1 | NCA1_CHICK | P13590 gallus galli |
| 47 | 120.5 | 5.2 | 3305 | 1 | APLP_MANSE | Q25490 manduca sex |
| 48 | 118.5 | 5.1 | 740 | 1 | PECI_PIG | Q95242 sus scrofa |
| 49 | 117.5 | 5.1 | 338 | 1 | LAMP_RAT | Q62813 ratu |
| 50 | 117.5 | 5.1 | 858 | 1 | NCA1_RAT | P13596 ratu |
| 51 | 117.5 | 5.1 | 1197 | 1 | CAML_BRARE | Q90478 brachydantio |
| 52 | 116 | 5.0 | 555 | 1 | C166_CARAU | Q90304 carassius a |
| 53 | 116 | 5.0 | 853 | 1 | NCA1_BOVIN | P13836 bos taurus |
| 54 | 115.5 | 5.0 | 879 | 1 | PRRP_RAT | Q62786 ratu |
| 55 | 115.5 | 5.0 | 1136 | 1 | TEI1_BOVIN | Q06805 bos taurus |
| 56 | 115.5 | 5.0 | 1302 | 1 | NRG_DROME | P20241 drosophilila |
| 57 | 114 | 4.9 | 257 | 1 | FCEA_HUMAN | P12319 homo sapien |
| 58 | 113.5 | 4.9 | 245 | 1 | FCBI_RAT | P12371 ratu |
| 59 | 113.5 | 4.9 | 338 | 1 | LAMP_CHICK | Q98919 gallus galli |
| 60 | 113.5 | 4.9 | 1277 | 1 | CAML_FUGRU | Q98902 tugu rubrip |
| 61 | 113.5 | 4.9 | 1849 | 1 | IG44_HAEIN | P45386 haemophilus |
| 62 | 113 | 4.9 | 1333 | 1 | VGR1_MOUSE | P35969 mus musculu |
| 63 | 112.5 | 4.8 | 359 | 1 | LACH_DROME | Q24372 drosophilila |
| 64 | 112 | 4.8 | 588 | 1 | C166_CHICK | P42292 gallus galli |
| 65 | 111.5 | 4.8 | 981 | 1 | ILIS_CERAE | Q29612 cercopithec |
| 66 | 111.5 | 4.8 | 981 | 1 | SCA4_RICE | Q8437 rickettsia |
| 67 | 110.5 | 4.8 | 879 | 1 | PRRP_MOUSE | Q94991 mus musculu |
| 68 | 110 | 4.7 | 777 | 1 | RNI1_RAT | Q64548 ratu |
| 69 | 110 | 4.7 | 1679 | 1 | GCC2_MOUSE | Q64548 ratu |
| 70 | 110 | 4.7 | 6632 | 1 | UN89_CAEBL | Q01761 caenorhabdi |
| 71 | 109 | 4.7 | 1367 | 1 | VGR2_MOUSE | P35918 mus musculu |
| 72 | 108 | 4.6 | 538 | 1 | PVR2_HUMAN | Q26262 homo sapien |
| 73 | 108 | 4.6 | 697 | 1 | STLA_HUMAN | Q96167 homo sapien |
| 74 | 107.5 | 4.6 | 1138 | 1 | TEI1_HUMAN | P35590 homo sapien |
| 75 | 107.5 | 4.6 | 1443 | 1 | NEOI_CHICK | Q90610 gallus galli |
| 76 | 107 | 4.6 | 514 | 1 | CSA_DICDI | P08796 dictyosteli |
| 77 | 107 | 4.6 | 521 | 1 | CEA1_MOUSE | P18809 mus musculu |
| 78 | 106.5 | 4.6 | 1022 | 1 | SCA4_RICPR | Q92448 rickettsia |
| 79 | 106 | 4.6 | 345 | 1 | OPCM_BOVIN | P18184 bos taurus |
| 80 | 105.5 | 4.5 | 916 | 1 | PERT_BORBR | Q03035 bordetella |
| 81 | 105.5 | 4.5 | 1150 | 1 | ZACA_HUMAN | Q06199 homo sapien |
| 82 | 105 | 4.5 | 327 | 1 | 112B_BOVIN | P46282 bos taurus |
| 83 | 105 | 4.5 | 345 | 1 | OPCM_HUMAN | Q14982 homo sapien |
| 84 | 105 | 4.5 | 1373 | 1 | RPOB_RICMA | Q91843 rickettsia |
| 85 | 104.5 | 4.5 | 515 | 1 | ENV_BIVAV | P25057 bovine leuk |
| 86 | 104.5 | 4.5 | 519 | 1 | ECTO_RAT | P16572 ratu |
| 87 | 104.5 | 4.5 | 1070 | 1 | PTK7_HUMAN | Q13308 homo sapien |
| 88 | 104 | 4.5 | 437 | 1 | E2F2_HUMAN | Q14209 homo sapien |
| 89 | 104 | 4.5 | 524 | 1 | BUTY_MOUSE | Q62556 mus musculu |
| 90 | 104 | 4.5 | 862 | 1 | CD22_MOUSE | P35329 mus musculu |
| 91 | 104 | 4.5 | 1284 | 1 | NRCA2_CHICK | P25331 gallus galli |
| 92 | 104 | 4.5 | 1348 | 1 | VGR2_COTJA | P52583 coturnix co |
| 93 | 104 | 4.5 | 1461 | 1 | NEOI_HUMAN | Q92859 homo sapien |
| 94 | 104 | 4.5 | 3210 | 1 | CENF_HUMAN | P19454 homo sapien |
| 95 | 103.5 | 4.5 | 108 | 1 | KY55_MOUSE | P01652 mus musculu |
| 96 | 103.5 | 4.5 | 129 | 1 | KY1W_HUMAN | P04431 homo sapien |
| 97 | 103.5 | 4.5 | 344 | 1 | NTRI_MOUSE | Q99910 mus musculu |
| 98 | 103.5 | 4.5 | 344 | 1 | NTRI_RAT | Q62718 ratu |
| 99 | 103.5 | 4.5 | 519 | 1 | GAG_SIVAT | P05892 samian immu |
| 100 | 103.5 | 4.5 | 738 | 1 | PECI_HUMAN | P14784 gallus galli |
| 101 | 103.5 | 4.5 | 1010 | 1 | CONT_CHICK | P14784 gallus galli |
| 102 | 103 | 4.4 | 327 | 1 | 112B_SHEEP | O02815 ovis aries |
| 103 | 103 | 4.4 | 562 | 1 | INVG_SALTY | P35672 salmoneilla |
| 104 | 103 | 4.4 | 1109 | 1 | POL_CAELV | P33459 caprine art |
| 105 | 103 | 4.4 | 1373 | 1 | RPOB_RICCN | Q91843 rickettsia |
| 106 | 102.5 | 4.4 | 515 | 1 | ENV_BIVAV | P25504 bovine leuk |

```

107 102.5 4.4 515 1 ENV BLVJ P03380 bovine leuk
108 102.5 4.4 521 1 LAG3 MOUSE O61790 mus musculu
109 102.5 4.4 1097 1 PGDR RAT O05010 rattus norv
110 102.5 4.4 1134 1 TIE1 MOUSE O06806 mus musculu
111 102 4.4 508 1 Y112_YEAST Q12303 saccharomyc
112 102 4.4 814 1 SLAI1 BACCA P49051 bacillus an
113 102 4.4 992 1 SCA4 RICRY Q9A163 rickettsia
114 102 4.4 1343 1 VGR2 RAT O08775 rattus norv
115 101.5 4.4 328 1 112B MACMU P48095 macaca mula
116 101.5 4.4 344 1 NTRI1 HUMAN O9P121 homo sapien
117 101 4.3 135 1 YOR6_ADEB1 P20748 avian adeno
118 101 4.3 1036 1 AXOI1 CHICK P28685 gallus gall
119 101 4.3 1541 1 IGAI1 HAEIN P42782 haemophilus
120 101 4.3 1694 1 SN MOUSE O62230 mus musculu
121 101 4.3 2029 1 LAR DROME P16621 drosophila
122 100.5 4.3 108 1 KVSQ_MOUSE P01650 mus musculu
123 100.5 4.3 597 1 STIL1_PANTR Q95110 pan troglod
124 100.5 4.3 822 1 FGRI1_HUMAN P11362 homo sapien
125 100 4.3 922 1 PERT1_BORPA P43328 bordetella

```

ALIGNMENTS

```

RESULT 1
CD4_HUMAN STANDARD; PRT; 458 AA.
AC P01730;

```

```

DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE T4/Leu-3).
GN GN
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85254948; PubMed=2990730;
RA Maddon P.J., Littman D.R., Godfrey M., Maddon D.E., Chess L.,
RA Axel R.;
RT "The isolation and nucleotide sequence of a cDNA encoding the T cell
RT surface protein T4: a new member of the immunoglobulin gene family.";
RL Cell 42:93-104(1985).
RN [2]
RP REVISION TO 26.
RX MEDLINE=89028665; PubMed=3263213;
RA Littman D.R., Maddon P.J., Axel R.;
RT "Corrected CD4 sequence.";
RL Cell 55:541-541(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96303695; PubMed=8723724;
RA Ananari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,
RA Malley T., Gibbs R.A.;
RT "A gene-rich cluster between the CD4 and triosephosphate isomerase
RT genes at human chromosome 12p13.";
RL Genome Res. 6:314-326(1996).
RN [4]
RP SEQUENCE FROM N.A. AND VARIANT TRP-265.
RX MEDLINE=91216786; PubMed=1708753;
RA Hodge T.W., Sasse D.R., McDougal J.S.;
RT "Humans with OKT8-epitope deficiency have a single nucleotide base
RT change in the CD4 gene, resulting in substitution of TRP240 for
RT ARG240.";
RL Hum. Immunol. 30:99-104(1991).
RN [5]
RP SEQUENCE FROM N.A.
RX TISSUE=Pancreas; PubMed=12477932;
RX MEDLINE=22386257;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

```

```

RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalka U., Smillius D.E.,
RA Schnerch A., Schein J.E., Jones S.T.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Eur. J. Immunol. 22:2973-2981(1992).
RN [7]
RP SEQUENCE OF 26-394.
RX MEDLINE=90078232; PubMed=2592374;
RA Carr S.A., Hemling M.E., Folsen-Wasserman G., Sweet R.W., Anumula K.,
RA Barr J.R., Huddleston M.J., Taylor P.;
RT "Protein and carbohydrate structural analysis of a recombinant
RT soluble CD4 receptor by mass spectrometry.";
RL J. Biol. Chem. 264:21286-21295(1989).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 26-208.
RX MEDLINE=91061881; PubMed=1701030;
RA Wang J., Yan Y., Garrett T.P., Liu J., Rodgers D.W., Garlick R.L.,
RA Tarr G.E., Husain Y., Reinherz E.L., Harrison S.C.;
RT "Atomic structure of a fragment of human CD4 containing two
RT immunoglobulin-like domains.";
RL Nature 348:411-418(1990).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 26-208.
RX MEDLINE=91061882; PubMed=2247146;
RA Ryu S.-E., Kwong P.D., Truneh A., Porter T.G., Arthos J.,
RA Rosenberg M., Dai X., Xiong N.-H., Axel R., Sweet R.W.,
RA Hendrickson W.A.;
RT "Crystal structure of an HIV-binding recombinant fragment of human
RT CD4.";
RL Nature 348:419-426(1990).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (3.9 ANGSTROMS) OF 26-388.
RX MEDLINE=97311402; PubMed=168119;
RA Wu H., Kwong P.D., Hendrickson W.A.;
RT "Bimetric association and segmental variability in the structure of
RT human CD4.";
RL Nature 387:527-530(1997).
RN [11]
RP PALMITOYLATION.
RX MEDLINE=92317088; PubMed=1618861;
RA Crise B., Rose J.K.;
RT "Identification of palmitoylation sites on CD4, the human
RT immunodeficiency virus receptor.";
RL J. Biol. Chem. 267:13593-13597(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

```

```

CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD4 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd4.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; M12807; AAA5572.1; -
DR EMBL; U47924; AAB51309.1; -
DR EMBL; M35160; AAL16069.1; -
DR EMBL; BC025782; AAB25782.1; -
DR PIR; A90872; RWH074.
DR PDB; 1CDH; 30-APR-94.
DR PDB; 1CDI; 30-APR-94.
DR PDB; 3CD4; 31-OCT-93.
DR PDB; 1CDJ; 01-APR-97.
DR PDB; 1CDU; 01-APR-97.
DR PDB; 1CDY; 01-APR-97.
DR PDB; 1WBR; 12-MAR-97.
DR PDB; 1WIO; 07-JUL-97.
DR PDB; 1WIP; 07-JUL-97.
DR PDB; 1WIO; 07-JUL-97.
DR PDB; 1G9M; 27-DEC-00.
DR PDB; 1G9N; 27-DEC-00.
DR PDB; 1GCI; 19-AUG-98.
DR PDB; 1JL4; 19-SEP-01.
DR GlycositeDB; P01730; -
DR Gene; HGNC:1678; CD4.
DR MIM; 166940; -
DR GO; GO:0042101; C:T-cell receptor complex; NAS.
DR GO; GO:0015026; F:coreceptor activity; NAS.
DR GO; GO:0015029; F:internalization receptor activity; TAS.
DR GO; GO:0042289; F:MHC class II protein binding; NAS.
DR GO; GO:0004888; F:transmembrane receptor activity; TAS.
DR GO; GO:0006955; F:immune response; NAS.
DR GO; GO:0009405; F:pathogenesis; TAS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . . NAS.
DR GO; GO:0030217; P:T-cell differentiation; NAS.
DR GO; GO:0045058; P:T-cell selection; NAS.
DR GO; GO:0007168; P:transmembrane receptor protein tyrosine kin. . . NAS.
DR InterPro; IPR000973; CD4_TcAg.
DR InterPro; IPR007110; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV_1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate; 3D-structure;
KW Polymorphism.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
FT LIPID 422 422
FT VARIANT 265 265

```

```

FT FT STRAND 27 32 /FTid=VAR_003906.
FT FT TURN 33 34
FT FT STRAND 37 39
FT FT TURN 44 45
FT FT STRAND 51 55
FT FT TURN 56 57

Query Match 87.4%; Score 2032; DB 1; Length 458;
Best Local Similarity 89.2%; Pred. No. 6.2e-132;
Matches 405; Conservative 2; Mismatches 17; Indels 30; Gaps 2;

OY 1 MNRGVPFRHLVLTQALLPAPATGKVVLAGKDDTVELCTASQKSIQPHWNSQIK 60
DB 1 MNRGVPFRHLVLTQALLPAPATGKVVLAGKDDTVELCTASQKSIQPHWNSQIK 60
OY 61 ILGQGSFLTKGPSKLNDRADRSRLWDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGQGSFLTKGPSKLNDRADRSRLWDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120
OY 121 LVFGLTANSDTHLQGSGLTITLSPGSSPSVOCSPRGKNIQGGKTLVSQLELDSG 180
DB 121 LVFGLTANSDTHLQGSGLTITLSPGSSPSVOCSPRGKNIQGGKTLVSQLELDSG 180
OY 181 TWTCVLOKOKVEFKIDIVLAFQKASITYKKEGSEVSPFLAFTVEKLTGSGELMW 240
DB 181 TWTCVLOKOKVEFKIDIVLAFQKASITYKKEGSEVSPFLAFTVEKLTGSGELMW 240
OY 241 QABRASSKSWITFDLNKKEVSVKRVTDPPKQNGKLPHLTLPOALPOYAGSGNITLA 300
DB 241 QABRASSKSWITFDLNKKEVSVKRVTDPPKQNGKLPHLTLPOALPOYAGSGNITLA 300
OY 301 LEATGKGLHDEVNLYVMRATQLOKNTLCEVWGSPSPKLMSTLKENAEAKSKREKVVW 360
DB 301 LEATGKGLHDEVNLYVMRATQLOKNTLCEVWGSPSPKLMSTLKENAEAKSKREKVVW 360
OY 361 LNPAGMOCILSPSGOVLSESNIKVLPWTSTPHRPAASALPAPPTGSALPDPTASALP 420
DB 361 LNPAGMOCILSPSGOVLSESNIKVLPWTSTPHRPAASALPAPPTGSALPDPTASALP 420
OY 421 DPPAASALPALAVISFLGLGV-ACVLARTR 453
DB 397 -----MALIVGAGVALLFLGIGIFPCVCRHR 425

RESULT 2
CD4_PANTR STANDARD; PRT; 458 AA.
AC P1604;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE T4/Leu-3).
GN CD4.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90182664; PubMed=2107024;
RA Camerini D.; Seed B.;
RT "A CD4 domain important for HIV-mediated syncytium formation lies
RT outside the virus binding site.";
RL Cell 60:747-754(1990).
RN [2]
RP SEQUENCE OF 26-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A.; Hirsch V.M.; Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human

```

```

RT Immunodeficiency virus.
RL Eur. J. Immunol. 22:2973-2981 (1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M31135; AAA35407.1; -.
DR EMBL, X73323; CAAS1749.1; -.
DR PIR, B32722; RMCZT4.
DR HSSP, P01730; 1MIO.
DR GO, GO:0042101; C:T-cell receptor complex; ISS.
DR GO, GO:0015026; F:coreceptor activity; ISS.
DR GO, GO:0042289; F:MHC class II protein binding; ISS.
DR GO, GO:0006955; P:immune response; ISS.
DR GO, GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO, GO:0030217; P:T-cell differentiation; ISS.
DR GO, GO:0045058; P:T-cell selection; ISS.
DR GO, GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR000973; CD4_TcRg.
DR InterPro: IPR007110; Ig_Like.
DR Pfam, PF00047; Ig_2.
DR PRINTS, PR00692; CD4TCANTIGEN.
DR SMART, SM00406; IGV_1.
DR PROSITE, PS50835; IG_Like, 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396 T-CELL SURFACE GLYCOPROTEIN CD4.
FT TRANSMEM 397 418 POTENTIAL.
FT DOMAIN 419 458 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 126 203 IG-LIKE V-TYPE.
FT DOMAIN 204 317 IG-LIKE C2-TYPE 1.
FT DOMAIN 318 374 IG-LIKE C2-TYPE 2.
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT DISULFID 41 109 BY SIMILARITY.
FT DISULFID 155 184 BY SIMILARITY.
FT DISULFID 328 370 BY SIMILARITY.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
FT CONFLICT 42 62 L -> N (IN REF. 2).
FT CONFLICT 62 62 L -> N (IN REF. 2).
FT CONFLICT 191 191 K -> N (IN REF. 2).
SQ SEQUENCE 458 AA; 51057 MW; ATGC3AC8A5257D3AD CMC64;
Query Match 86.4%; Score 2008; DB 1; Length 458;
Best Local Similarity 88.3%; Pred. No. 2,7e-130;
Matches 401; Conservative 2; Mismatches 21; Indels 30; Gaps 2;

```

```

Db 121 LVFGLTANSDTHLLQGSGLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELQDSG 180
Qy 181 TWTCVTLQNKQKVEFKIDIVLAFQKASSIVYKKEGQVFFSPPLATFTVEKLTGSGELMW 240
Db 181 TWTCVTLQNKQKVEFKIDIVLAFQKASSIVYKKEGQVFFSPPLATFTVEKLTGSGELMW 240
Qy 241 QAERASSSKSWITFDLKNKEVSVKRYTODPKLQMGKKLPHLTLPOALPOYAGSGNLTLA 300
Db 241 QAERASSSKSWITFDLKNKEVSVKRYTODPKLQMGKKLPHLTLPOALPOYAGSGNLTLA 300
Qy 301 LEAKTGLHGEVNLVVMRAQLQKNTLCEYWGFTSPPLMLSLKLNKAAYSKREKVMV 360
Db 301 LEAKTGLHGEVNLVVMRAQLQKNTLCEYWGFTSPPLMLSLKLNKAAYSKREKVMV 360
Qy 361 LNEBAGMOCQLSDSGQVLESNIKYLPTWSTPVHPPASALPAPPTGSALPDQTASALP 420
Db 361 LNEBAGMOCQLSDSGQVLESNIKYLPTWSTPVHPPASALPAPPTGSALPDQTASALP 420
Qy 421 DPPAASALPALAVISFLGLGV-ACVLARTR 453
Db 397 -----MALIVLGVAGLLFTLGIFFCVRCRR 425

RESULT 3
ID CD4_MACFU STANDARD; PRT; 458 AA.
AC P79184;
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
GN CD4.
OS Macaca fuscata fuscata (Japanese macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_Taxid=9543;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto O., Tatemui M.;
RL Submitted (FEBS-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, D63348; BAA09672.1; -.
DR HSSP, P01730; 1MIO.
DR GO, GO:0042101; C:T-cell receptor complex; ISS.
DR GO, GO:0015026; F:coreceptor activity; ISS.
DR GO, GO:0042289; F:MHC class II protein binding; ISS.
DR GO, GO:0006955; P:immune response; ISS.
DR GO, GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO, GO:0030217; P:T-cell differentiation; ISS.
DR GO, GO:0045058; P:T-cell selection; ISS.
DR InterPro: IPR007110; Ig_Like.
DR InterPro: IPR003596; Ig_v.
DR Pfam, PF00047; Ig_2.
DR PRINTS, PR00406; IGV_1.
DR SMART, SM00406; IGV_1.

```

```

DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 458
FT TRANSMEM 26 396
FT DOMAIN 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT CAROXYD 318 374
FT CAROXYD 42 42
FT CAROXYD 296 325
FT CAROXYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
SQ SEQUENCE 458 AA; 50828 MW; 76B3E7E081B5535 CRC64;

Query Match 79.9%; Score 1858; DB 1; Length 458;
Best Local Similarity 81.3%; Pred. No. 5.1e-120;
Matches 369; Conservative 18; Mismatches 37; Indels 30; Gaps 2;

QY 1 MNRGVPFRLHLLVLTALIPATQGNKYVGGKGTVEITCTASQKSIQFMKNSNOIK 60
DB 1 MNRGVPFRLHLLVLTALIPATQGNKYVGGKGTVEITCTASQKSIQFMKNSNOIK 60
QY 61 ILGNQSFLLTKGPSKLNDRADRSRLMDQGNPLIKLKLTEDSDTYICEVDOKEEYOL 120
DB 61 ILGNQSFLLTKGPSKLNDRADRSRLMDQGNPLIKLKLTEDSDTYICEVDOKEEYOL 120
QY 121 LVFGLTANSDFHLLGQSLITLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDFHLLGQSLITLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 121 LVFGLTANSDFHLLGQSLITLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDFHLLGQSLITLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLOQKQKVEKIDIVLAFOKASSIVYKKEGEQVSEFPLAFVEKLTSGGELMW 240
DB 181 TWTCTVLOQKQKVEKIDIVLAFOKASSIVYKKEGEQVSEFPLAFVEKLTSGGELMW 240
QY 181 TWTCTVLOQKQKVEKIDIVLAFOKASSIVYKKEGEQVSEFPLAFVEKLTSGGELMW 240
DB 181 TWTCTVLOQKQKVEKIDIVLAFOKASSIVYKKEGEQVSEFPLAFVEKLTSGGELMW 240
QY 241 QAEKSSSSKSWITTPOLKKEVSVKVTOPKLOMGKKPLHLTLPOALPOVAGSNNLTILA 300
DB 241 QAEKSSSSKSWITTPOLKKEVSVKVTOPKLOMGKKPLHLTLPOALPOVAGSNNLTILA 300
QY 301 LEAKTGKLEHGVNVLVVRATQLOKNTLCEVWGPTSPKLMLSIKLENKAKYSKREKPYVW 360
DB 301 LEAKTGKLEHGVNVLVVRATQLOKNTLCEVWGPTSPKLMLSIKLENKAKYSKREKPYVW 360
QY 301 LEAKTGKLEHGVNVLVVRATQLOKNTLCEVWGPTSPKLMLSIKLENKAKYSKREKPYVW 360
DB 301 LEAKTGKLEHGVNVLVVRATQLOKNTLCEVWGPTSPKLMLSIKLENKAKYSKREKPYVW 360
QY 361 LNPEAGMOCCLISDSGOVLLESNIVYLPWSTPPVHPRASALPAPPTGSALPDPTASALP 420
DB 361 LNPEAGMOCCLISDSGOVLLESNIVYLPWSTPPVHPRASALPAPPTGSALPDPTASALP 420
QY 421 DEPASALPALAVISFLGLGLGV-ACYLARTR 453
DB 421 DEPASALPALAVISFLGLGLGV-ACYLARTR 453
QY 397 -----MALIVLGAVAGLLFTGLGFPCVCRHR 425
DB 397 -----MALIVLGAVAGLLFTGLGFPCVCRHR 425

RESULT 4
CD4_MACMU STANDARD; PRT; 458 AA.
AC P16T03; Q29617;
DT 01-APR-1990 (Rel. 14, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell) surface antigen
DE T4/Leu-3).
GN CD4.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopitheciinae; Macaca.

```

```

OX NCBI_TaxID=9544;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=90182664; PubMed=2107024;
RA Cameron D.; Seed B.;
RT "A CD4 domain important for HIV-mediated syncytium formation lies
RL outside the virus binding site.";
RN Cell 60:747-754(1990).
RN [2]
RN SEQUENCE FROM N.A.
RN TISSUE=Thymocytes;
RA Hashimoto O.; Tatsumi M.;
RT "Molecular cloning and expression of macaque CD4s.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 28-424 FROM N.A.
RN TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A.; Hirsch V.M.; Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RL cellular receptor for simian immunodeficiency virus/human
RL immunodeficiency virus.";
RL Eur. J. Immunol. 22:2973-2981(1992).
RN [4]
RN SEQUENCE OF 107-192 FROM N.A.
RX MEDLINE=98320644; PubMed=9656488;
RA Harris E.B.; DiBotto T.R.;
RT "Nuclear gene trees and the phylogenetic relationships of the
RL mangabey (Primates: Papionini).";
RL Mol. Biol. Evol. 15:892-900(1998).
RN [5]
RN FUNCTION: Accessory protein for MHC class-II antigen/T-cell
RN receptor interaction. May regulate T-cell activation.
RN - SUBUNIT: Associates with p56-lck (By similarity).
RN - SUBCELLULAR LOCATION: Type I membrane protein.
RN - SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
RN - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
RN -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; M31134; AAA36838.1; -
DR EMBL; D63347; BAA09671.1; -
DR EMBL; X73326; CAA51752.1; -
DR EMBL; AF057385; AAC25129.1; -
DR HSPB; P01730; IMB.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; P:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003596; IG_v.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV_1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125

```

```

FT DOMAIN 126 203 IG-LIKE C2-TYPE 1.
FT DOMAIN 204 317 IG-LIKE C2-TYPE 2.
FT DOMAIN 318 374 IG-LIKE C2-TYPE 3.
FT CARBOHYD 296 326 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT DISULFID 41 109 BY SIMILARITY.
FT DISULFID 155 184 BY SIMILARITY.
FT DISULFID 328 370 BY SIMILARITY.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
FT CONFLICT 42 42 N -> T (IN REF. 1).
FT CONFLICT 62 62 L -> S (IN REF. 3).
FT CONFLICT 67 67 L -> S (IN REF. 2).
FT CONFLICT 169 169 I -> L (IN REF. 2).
FT CONFLICT 191 191 K -> N (IN REF. 3).
FT CONFLICT 248 248 S -> P (IN REF. 2).
FT CONFLICT 265 265 R -> Q (IN REF. 3).
FT CONFLICT 349 349 A -> T (IN REF. 2).
SQ SEQUENCE 458 AA; 50884 MW; 8BB80339FAFC808 CRC64;

```

Query Match 79.9%; Score 1857; DB 1; Length 458;

Best Local Similarity 81.3%; Pred. No. 6e-120; Mismatches 37; Indels 30; Gaps 2;

Matches 369; Conservative 18; Mismatches 37; Indels 30; Gaps 2;

```

QY 1 MNRGVPFRHLVLVQLALPAAVQGNKVVLGKGGDTVELTCTASOKKSIQFHMKNNSQIK 60
DB 1 MNRGVPFRHLVLVQLALPAAVQGNKVVLGKGGDTVELTCTASOKKSIQFHMKNNSQIK 60
QY 61 ILNGQSFLLTKGSKLNDRAKRSRLMDQGNFLLIKLKIEDSDTYICEVNDKEVQL 120
DB 61 ILNGQSFLLTKGSKLNDRAKRSRLMDQGNFLLIKLKIEDSDTYICEVNDKEVQL 120
QY 121 LVFGLTANSPTHLLOGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDSG 180
DB 121 LVFGLTANSPTHLLOGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDSG 180
QY 121 LVFGLTANSPTHLLOGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDSG 180
DB 121 LVFGLTANSPTHLLOGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDSG 180
QY 181 TWTCVLOMOKKVEFKIDIVLAFOKASSIYKKEGEVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWTCVLOMOKKVEFKIDIVLAFOKASSIYKKEGEVEFSPLAFTVEKLTGSGELMW 240
QY 181 TWTCVLOMOKKVEFKIDIVLAFOKASSIYKKEGEVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWTCVLOMOKKVEFKIDIVLAFOKASSIYKKEGEVEFSPLAFTVEKLTGSGELMW 240
QY 241 QAEBSASSKSWITPDLKKEVSVKRVODPKLOMGKTLPLTLTPQALPOVAGSGNLTLA 300
DB 241 QAEBSASSKSWITPDLKKEVSVKRVODPKLOMGKTLPLTLTPQALPOVAGSGNLTLA 300
QY 241 QAEBSASSKSWITPDLKKEVSVKRVODPKLOMGKTLPLTLTPQALPOVAGSGNLTLA 300
DB 241 QAEBSASSKSWITPDLKKEVSVKRVODPKLOMGKTLPLTLTPQALPOVAGSGNLTLA 300
QY 301 LEAKTGKLGHOENVLVVWRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKKEKPVW 360
DB 301 LEAKTGKLGHOENVLVVWRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKKEKPVW 360
QY 301 LEAKTGKLGHOENVLVVWRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKKEKPVW 360
DB 301 LEAKTGKLGHOENVLVVWRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKKEKPVW 360
QY 361 LNPEAGMOCQLSDSGOVLLESNITKVLPTWSTPVHPRASALPAPPTGSLPDPQTASALP 420
DB 361 LNPEAGMOCQLSDSGOVLLESNITKVLPTWSTPVHPRASALPAPPTGSLPDPQTASALP 420
QY 421 DPASASLPALAVISFLGLGLGV-ACVLARTR 453
DB 421 DPASASLPALAVISFLGLGLGV-ACVLARTR 453
QY 397 -----MALIVGVAGVGLLFTGLGIFFCVRCRHR 425
DB 397 -----MALIVGVAGVGLLFTGLGIFFCVRCRHR 425

```

RESULT 5

CD4_MACFA STANDARD; PRT; 458 AA.

```

AC P79185;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
   T4/Leu-3).
GN CD4.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Thymocytes;
RA Tatsumi M., Yabe M., Yamada Y.K.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -!- SUBUNIT: Associates with p56-lck (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

DR EMBL; D63349; BAA09673.1; -.
DR HSSP; P01730; IWR.
DR GO; GO:0042101; C-T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro; IPR000973; CD4-TCRg.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTGEN.
DR SMART; SM00406; IGV_1.
DR PROSITE; PS50835; IG_LIKE_1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 418
FT TRANSLEM 397 418
FT DOMAIN 419 458
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 42 42
FT CARBOHYD 296 326
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
FT LIPID 422 422
SQ SEQUENCE 458 AA; 50872 MW; 9105479FB5C56FF7 CRC64;

```

Query Match 79.7%; Score 1854; DB 1; Length 458;

Best Local Similarity 81.3%; Pred. No. 9.6e-120; Mismatches 38; Indels 30; Gaps 2;

Matches 369; Conservative 17; Mismatches 38; Indels 30; Gaps 2;

```

QY 1 MNRGVPFRHLVLVQLALPAAVQGNKVVLGKGGDTVELTCTASOKKSIQFHMKNNSQIK 60
DB 1 MNRGVPFRHLVLVQLALPAAVQGNKVVLGKGGDTVELTCTASOKKSIQFHMKNNSQIK 60
QY 61 ILNGQSFLLTKGSKLNDRAKRSRLMDQGNFLLIKLKIEDSDTYICEVNDKEVQL 120
DB 61 ILNGQSFLLTKGSKLNDRAKRSRLMDQGNFLLIKLKIEDSDTYICEVNDKEVQL 120
QY 121 LVFGLTANSPTHLLOGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDSG 180
DB 121 LVFGLTANSPTHLLOGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDSG 180

```

| | | | |
|----|-----|--|-----|
| Qy | 18 | TMTCTVLONOKKKEPKIDIVLAFQKASSIVYKKEGEVESPFLAIFVEKLTGSGSLM | 240 |
| Db | 181 | TMTCTVSQDDKVTEFKIDIVLAFQKASTYVKKEGEEVESPFLATLEKLTGSSGLM | 240 |
| Qy | 241 | QAERASSSKSWITFDLNKNKEVSVKRYTODPKLQMGKKPLHLTLPOALPYAGSGNLTLLA | 300 |
| Db | 241 | QAERASSSKSWITFDLNKNKEVSVKRYTODPKLQMGKKPLHLTLPOALPYAGSGNLTLLA | 300 |
| Qy | 301 | LEAKTKGLHQEVNLVMRATLOLKULTCEWMEPTSPKMLSLIKLENKAISKREKRVNV | 360 |
| Db | 301 | LEAKTKGLHQEVNLVMRATLOEQENLTCEWMEPTSPKLTLSLKENKGTITYSKAKAVNW | 360 |
| Qy | 361 | LNEBAGMWQCILSDSGOVLNESNIKYLPTWSIPVPRASALPAPPTGSALPDPTASALP | 420 |
| Db | 361 | LNEBAGMWQCILSDSGOVLNESNIKVPMPPIPVGP----- | 366 |
| Qy | 421 | DPPRASALPALNAVISFLLGLGLGV-ACCLTAATR | 453 |
| Db | 397 | -----MALIVLGAGVALLFTGTGGIETFCVCRCRRH | 425 |

```

RESULT 6
ID4_MACNE          STANDARD;          PRT;          458 AA.
AC      008340; P79196;
DT      01-FEB-1995 (Rel. 31, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE      T4/Lew-3).
GN      CD4.
OS      Macaca nemestrina (Pig-tailed macaque).
OC      Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC      Cercopithecinae; Macaca.
OX      NCBI_TaxID=9545;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Hashimoto O., Tateumi M.;
RL      Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
RC      [2]
RP      SEQUENCE OF 28-424 FROM N.A.
RT      TISSUE=Blood;
RX      MEDLINE=93049640; PubMed=1425921;
RA      Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT      "Cloning and sequences of primate CD4 molecules: diversity of the
RT      cellular receptor for simian immunodeficiency virus/human
RT      immunodeficiency virus.";
RL      Eur. J. Immunol. 22:2973-2981(1992).
CC      -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC      receptor interaction. May regulate T-cell activation.
CC      -1- SUBUNIT: Associates with p56-lck (By similarity).
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC      -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC      -----
CC      This SWISS-PROT entry is copyrighted. It is produced through a collaboratio
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on it
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce
CC      or send an email to license@isb.slb.ch).
CC      -----
DR      EMBL; D63346; BAA09670.1; -;
DR      EMBL; X73325; CAAS1751.1; -;
DR      HSSP; P01730; 1WBK.
DR      GO; GO:0042101; G:T-cell receptor complex; ISS.
DR      GO; GO:0015026; F:coreceptor activity; ISS.
DR      GO; GO:0042289; F:MHC class II protein binding; ISS.
DR      GO; GO:0006955; P:immune response; ISS.
DR      GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
DR      GO; GO:0030217; P:T-cell differentiation; ISS.

```

| | | | |
|-----|---|---|--------------------------|
| DR | GO: 0045058 | P: T-cell selection | ISS. |
| DR | GO: 0007169 | P: Transmembrane receptor protein tyrosine kin. . . | ISS. |
| DR | InterPro: IPR000373 | CD4_TGAg. | |
| DR | InterPro: IPR007110 | IG-like. | |
| DR | Pfam: PF00047 | 1g_2. | |
| DR | PRINTS: PR00692 | CD4TCANTIGEN. | |
| DR | SMART: SM00406 | IGV: 1. | |
| DR | PROSITE: PS00835 | IG-LIKE; 1. | |
| KW | Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell; | | |
| KW | Immune response; Repeat; Signal; Lipoprotein; Palmitate. | | |
| FT | SIGNAL | 1 | 25 |
| FT | CHAIN | 26 | 458 |
| FT | DOMAIN | 26 | 396 |
| FT | TRANSMEM | 397 | 418 |
| FT | DOMAIN | 419 | 458 |
| FT | DOMAIN | 26 | 125 |
| FT | DOMAIN | 126 | 203 |
| FT | DOMAIN | 204 | 317 |
| FT | DOMAIN | 318 | 374 |
| FT | CARBOHYD | 42 | 42 |
| FT | CARBOHYD | 296 | 296 |
| FT | CARBOHYD | 325 | 325 |
| FT | DISULFID | 41 | 109 |
| FT | DISULFID | 155 | 184 |
| FT | DISULFID | 328 | 370 |
| FT | LIPID | 419 | 419 |
| FT | LIPID | 422 | 422 |
| FT | CONFLICT | 57 | 57 |
| FT | CONFLICT | 91 | 91 |
| FT | CONFLICT | 105 | 105 |
| FT | CONFLICT | 113 | 113 |
| FT | CONFLICT | 302 | 302 |
| FT | CONFLICT | 349 | 349 |
| SEQ | SEQUENCE | 458 AA; 50905 MM; | 751A9BA2C8B3EBE16 CRC64; |

| | | | | |
|-----------------------|------------------|---------------------|------------|-------------|
| Query Match | 79.5% | Score 1849; | DB 1; | Length 458; |
| Best Local Similarity | 80.8%; | Pred. No. 2.1e-119; | | |
| Matches 367; | Conservative 20; | Mismatches 37; | Indels 30; | Gaps 2 |

```

OY 1 MNRBVPFRHLLVYQALMLPAAQTGNVYLGGKDDYELTCTASOKS10PHMKNSNOIK 60
Db 1 MNRG1PFRHLLVYQALMLPAAVTGQKVKVLGKKDDYELTCTASQKKNTPHMKNSDQIK 60
OY 61 ILAGQGSFLTGSPSKLNDRADRSRLMDQGNFLIKLKLKIEDSDYIVCEYDQKEEYQL 120
Db 61 ILG1QGSFLTGSPSKLSDRADRSKSLMDQGFMSIMIKLKLIEDSNTYIVCEVENEKEEVEL 120
OY 121 LVFBLTANSDBHLIQQGSLTLTLESPPGSSPVOCRSFRGNKTIQGGKTLVSQLELQDSG 180
Db 121 LVFGLTANSDBHLIEGOSLTLTLESPPSSPVCSRSPGKGNIOGGRTLSPOLEROQSG 180
OY 181 TWTCTV1ONOKKVEFKDIDIVLAFOKASSIYYKKKEGEVRSFPLAFVLEKLTGSGELMM 240
Db 181 TWTCTSQDQKTVFEPKIDIVLAFOKASSIYYKKKEGEVRSFPLAFVLEKLTGSGELMM 240
OY 241 QABRASSKSMITFDLKNKEVSVKRVTOPDLQMGKCLPLHLTLPOALPOYAGSGNLTLLA 300
Db 241 QABRASSKSMITFDLKNKEVSVKRVTOPDLQMGKCLPLHLTLPOALPOYAGSGNLTLLA 300
OY 301 LEATTGKLHGVN1LVNMRATQLOKNLCEVWGGPSPFLKMLSLKLENKEAKSKREKPVWV 360
Db 301 LEATTGKLHGVN1LVNMRATQFOENLTCEVWGGPSPFLKMLSLKLENKGTVYSKAKAVWV 360
OY 361 LNPBAGMMQCLSDSGQVLTESN1KVLPTWSTPAPHPRASALPAPPTGSLALPDQTASALP 420
Db 361 LNPBAGMMQCLSDSGQVLTESN1KVLPTWSTPAPHPRASALPAPPTGSLALPDQTASALP 420
OY 421 DDPASALPAA1AVISFLIGLGV-ACVTLARTR 453
Db 397 -----MALIVGGVAGLLPFGTGLIFPCVRCRHR 425

```

```

RESULT 7
CD4_CERAB STANDARD: PRT: 458 AA.
AC 008338; 002805; 07593; 028217;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
  T4/Leu-3).
EN CD4.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto O., Tatsumi M.,
RT "Molecular cloning and expression of african green monkey CD4.",
RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93043640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.,
RT "Cloning and sequences of primate CD4 molecules: diversity of the
  cellular receptor for simian immunodeficiency virus/human
  immunodeficiency virus.",
RL Eur. J. Immunol. 22:2973-2981(1992).
RN [3]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Peripheral blood;
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Mueller-Trutwin M.C., Diop O., Hansen J., Machiot C.,
  Corbet S., Barre-Sinoussi F., Allan J.S.,
RT "Relation between phylogeny of African green monkey CD4 genes and
  their respective simian immunodeficiency virus genes.",
RL J. Med. Primatol. 26:120-128(1997).
RN [4]
RP SEQUENCE OF 107-192 FROM N.A.
RX MEDLINE=98320644; PubMed=9656488;
RA Harris E.B., DiCorleil T.R.,
RT "Nuclear gene trees and the phylogenetic relationships of the
  mangabeys (Primates: Papionini).",
RL Mol. Biol. Evol. 15:892-900(1998).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
  receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D86589; BAAL1332.1; -
DR EMBL; X73322; CA51748.1; -
DR EMBL; AF001226; AAB60873.1; -
DR EMBL; AF001228; AAB60875.1; -
DR EMBL; AF057380; AAC25124.1; -
DR HSPB; P01730; IWO.
DR GO; GO:004210; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.

```

```

DR GO; GO:0007169; P:Transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR000973; CD4_TcAg.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KM Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
  Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSSEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 42 42
FT CARBOHYD 281 281
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
FT LIPID 422 422
FT CONFLICT 46 46
FT CONFLICT 59 59
FT CONFLICT 115 115
FT CONFLICT 165 165
FT CONFLICT 200 200
FT CONFLICT 227 227
FT CONFLICT 271 271
FT CONFLICT 281 281
SQ SEQUENCE 458 AA; 51158 MW; FCS23D2EDD1F72B7 CRC64;

Query Match 78.8%; Score 1832; DB 1; Length 458;
Best Local Similarity 80.4%; Pred. No. 3,1e-118;
Matches 365; Conservative 18; Mismatches 41; Indels 30; Gaps 2;

QY 1 MNRGVPFRHLVLTQALPAATQGNKRVVLGKKGDTVELTCTASQKKSIGFHNKSNQIK 60
DB 1 MNRGVPFRHLVLTQALPAATQGNKRVVLGKKGDTVELTCTASQKKSIGFHNKSNQIK 60
QY 61 ILNGQSFLLTKGSKLANDRSDRSIMDQGNFLLIKNLKIDSDTYICVEBQKEVQL 120
DB 61 ILNGQSFLLTKGSKLANDRSDRSIMDQGNFLLIKNLKIDSDTYICVEBQKEVQL 120
QY 61 ILKQGSFLLTKGSKLRDRIDSRKSLMDQCFSMIIKNLKIETSEYICVENKKEVEL 120
DB 61 ILKQGSFLLTKGSKLRDRIDSRKSLMDQCFSMIIKNLKIETSEYICVENKKEVEL 120
QY 121 LVFGLTANSDTHLLQGOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLQNGKVEFKIDIVLAFOKASSIVYKKEGBOYFSPPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLQNGKVEFKIDIVLAFOKASSIVYKKEGBOYFSPPLAFTVEKLTGSGELMW 240
QY 181 TWTCTVSDQNTVEFKIDIVLAFOKASSIVYKKEGBOYFSPPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVSDQNTVEFKIDIVLAFOKASSIVYKKEGBOYFSPPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPHLTLPLPOLPYASSGNITLA 300
DB 241 QAERASSSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPHLTLPLPOLPYASSGNITLA 300
QY 301 LEAKTGKHOEVLLVYMRATQLOKNLTCEYWGTSPLMLSLLENBEAVSREKPVVW 360
DB 301 LEAKTGKHOEVLLVYMRATQLOKNLTCEYWGTSPLMLSLLENBEAVSREKPVVW 360
QY 361 LNEBAGMOCILSDSQVLLIESNIKYLPWTSTFVHPRASALPAPPTGSALPDQRTASALP 420
DB 361 LNEBAGMOCILSDSQVLLIESNIKYLPWTSTFVHPRASALPAPPTGSALPDQRTASALP 420
QY 421 DPAASALPAAALVIFSLGLGLGV-ACVLARTR 453
DB 421 DPAASALPAAALVIFSLGLGLGV-ACVLARTR 453
QY 397 -----MALIVLGGVAGLLFTGLGIFFCVCRHR 425
DB 397 -----MALIVLGGVAGLLFTGLGIFFCVCRHR 425

```

```

RESULT 8
CD4_ERYPA STANDARD: PRT: 397 AA.
ID CD4_ERYPA
AC Q08339;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Lew-3)
   (Fragment).
GN CD4.
OS Erythrocybus patas (Red guenon) (Husar).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Erythrocybus.
OC NCBI_TaxId=9538;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.
RT Bur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X73324; CAA51750.1; -.
CC HSP; P01730; IMIO.
DR GO: GO:0042101; C:T-cell receptor complex; ISS.
DR GO: GO:0015026; F:coreceptor activity; ISS.
DR GO: GO:0042289; F:MHC class II protein binding; ISS.
DR GO: GO:0006955; P:immune response; ISS.
DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO: GO:0030217; P:T-cell differentiation; ISS.
DR GO: GO:0045058; P:T-cell selection; ISS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR000973; CD4_TcAg.
DR InterPro: IPR007110; Ig_V.
DR Pfam: PF00047; Ig_2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SMO0406; IGV_1.
DR PROSITE: PS50835; IG_LIKE_1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Lipoprotein; Palmitate.
FT NON TER 1
FT DOMAIN <1 369 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 370 391 POTENTIAL.
FT DOMAIN 392 397 CYTOPLASMIC (POTENTIAL).
FT DOMAIN <1 98 IG-LIKE V-TYPE.
FT DOMAIN 99 176 IG-LIKE C2-TYPE 1.
FT DOMAIN 177 290 IG-LIKE C2-TYPE 2.
FT DOMAIN 291 347 IG-LIKE C2-TYPE 3.
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT DISULFID 14 82 BY SIMILARITY.
FT DISULFID 128 157 BY SIMILARITY.
FT DISULFID 301 343 BY SIMILARITY.

```

```

FT LIPID 392 392 S-palmitoyl cysteine (By similarity).
FT LIPID 395 395 S-palmitoyl cysteine (By similarity).
FT NON TER 397 397
SQ SEQUENCE 397 AA; 44081 MW; 678879797A6B7BA4F CRC64;
Query Match 74.3%; Score 1728; DB 1; Length 397;
Best Local Similarity 80.5%; Pred. No. 3,4e-11;
Matches 342; Conservative 16; Mismatches 37; Indels 30; Gaps 2;
QY 28 VVLGKGGTVLTLCTASGKSIQFHWKNSNQIKLIGNGSFLTKPSTLNPADRSRLM 87
DB 1 VVLGKGGTVLTLCTASGKSIQFHWKNSNQIKLIGNGSFLTKPSTLNPADRSRLM 60
QY 88 DQGNPLILIKMLKIEDSTY;CEVEDKEEYQVLLFGLTASDPTLLGGSTLTLRESPP 147
DB 61 DQGCSTMIITKMLKIEDSTY;CEVEDKEEYVLLFGLTASDPTLLGGSTLTLRESPP 120
QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLQNKKEFKIDIVLAFQKA 207
DB 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLQNKKEFKIDIVLAFQKA 180
QY 208 SSIVYKKEGEYERSFPLAFVEKLTGSGELMWQAEARSSSKSMTTFLPKNKEVSVKYT 267
DB 191 SSIVYKKEGEYERSFPLAFVEKLTGSGELMWQAEARSSSKSMTTFLPKNKEVSVKYT 240
QY 268 QDPKLQMGKKLPLHLTLPOALPOYAGSGLTLLEAKTGKXHOEVNLYVMRATOLQKMLT 327
DB 241 QDPKLQMGKKLPLHLTLPOALPHYAGSGLTLLEAKTGKXHOEVNLYVMRATOLQKMLT 300
QY 328 CEVWGPTSPKMLSLIKENKEAKVSKREKPYVVLNPEAGMQLLSDSGOVLLESNIVLT 387
DB 301 CEVWGPTSPKMLSLIKENKEAKVSKREKPYVVLNPEAGMQLLSDSGOVLLESNIVLT 360
QY 368 PTVSPVHPRASALPAPPTGSALPDPPTASALPDPASALPALALVSLFLGLGLSV-A 446
DB 361 PTVSPVHPRASALPAPPTGSALPDPPTASALPDPASALPALALVSLFLGLGLSV-A 446
QY 447 CYLAR 451
DB 392 CVRCR 396
RESULT 9
CD4_CERTO STANDARD: PRT: 397 AA.
ID CD4_CERTO
AC Q08336;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Lew-3)
   (Fragment).
GN CD4.
OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OC NCBI_TaxId=9531;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.
RT Bur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X73327; CAAS1754.1; -
 DR EMBL: X73327; CAAS1753.1; -
 DR HSPB; P01730; 1M7O.
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.
 DR GO; GO:0015026; F:coreceptor activity; ISS.
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.
 DR GO; GO:0006955; P:immune response; ISS.
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
 DR GO; GO:0030217; P:T-cell differentiation; ISS.
 DR GO; GO:0045058; P:T-cell selection; ISS.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
 DR InterPro: IPR000973; CD4 TCAG.
 DR InterPro: IPR007110; IG-like.
 DR InterPro: IPR003596; IG_v.
 DR Pfam; PF00047; IG; 2.
 DR PRINTS; PR00692; CD4TCANTIGEN.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
 KW Immune response; Repeat; Lipoprotein; Palmitate.
 FT NON_TER 1
 FT DOMAIN 1 369
 FT TRANSMM 370 391
 FT DOMAIN 392 >397
 FT DOMAIN <1 98
 FT DOMAIN 99 176
 FT DOMAIN 177 290
 FT DOMAIN 291 347
 FT CARBOHYD 15
 FT CARBOHYD 269
 FT CARBOHYD 298
 FT DISULFID 14
 FT DISULFID 128
 FT DISULFID 301
 FT LIPID 392 395
 FT LIPID 395 392
 FT VARIANT 20 20
 FT VARIANT 43 43
 FT VARIANT 86 86
 FT VARIANT 96 96
 FT VARIANT 173 173
 FT VARIANT 316 316
 FT NON_TER 397
 SQ SEQUENCE 397 AA; 43926 MW; 8660B636D2DB38A7 CRC64;

Query Match 73.9%; Score 1718; DB 1; Length 397;
 Best Local Similarity 80.2%; Pred. No. 1.6e-110;
 Matches 341; Conservative 19; Mismatches 35; Indels 30; Gaps 2;

QY 28 VVLGGKGTVELTCTASQKSKSQFHWKNSKQIKLGNQGSFLTKGSSKLSRADSRSKLW 87
 DB 1 VVLGGKGTVELTCTASQKSKSQFHWKNSKQIKLGNQGSFLTKGSSKLSRADSRSKLW 60
 QY 88 DQGNFLLIKNLKLTEDSDYICEDVQKEEVOLVFGLTANSDFHLQGSLLTLTLESP 147
 DB 61 DQGCSTMIKLIKLTEDSETYICEVKNKEVELVFGLTANSDFHLQGSLLTLTLESP 120
 QY 148 GSSPEVQCRSPKNGKIQGKGLTSVSELELQDSGWTCTVLQNKQKVEFKIDIVLAFOKA 207
 DB 121 GSSPEVKCRSPKNGKIQGKGLTSVSELELQDSGWTCTVLQNKQKVEFKIDIVLAFOKA 180
 QY 208 SSIYVKKGEQVEFSFLAFVTEKLTGSGELMWQKERRASSKWTPLTKKKEVSVKRV 267
 DB 181 SSIYVKKGEQVEFSFLAFVTEKLTGSGELMWQKERRASSKWTPLTKKKEVSVKRV 240

QY 268 QDPFLQWGGKLLPHLTLPQALPOVAGSGNTLALFAKTKGLHDEVNLVWKRATOLQKRLT 327
 DB 241 QDPFLQWGGKLLPHLTLPQALPOVAGSGNTLALFAKTKGLHDEVNLVWKRATOLQKRLT 300
 QY 328 CEVWGPPSPKLMILSKLENKAVSKREKPVWVNLNPAQMOCILSDSGOVLESNIKVL 387
 DB 301 CEVWGPPSPKLTSLRLNKKATVSKQAKAVWVNLNPAQMOCILSDSGOVLESNIKVL 360
 QY 388 PTWSTPVHPASALPAPPTGSALPDPQTASALDPAPASALPALANVLSFLGLGLGV-A 446
 DB 361 PTWSTPVHP-----MALIVGVAGLLFTGLGIFP 391
 QY 447 CVLAR 451
 DB 392 CVRCR 396

RESULT 10
 CD4_SAISC STANDARD; PRT; 457 AA.
 ID CD4_SAISC
 AC Q29037;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
 DE T4/Leu-3).
 GN CD4.
 OS Saimiri sciureus (Common squirrel monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
 OX NCBI_TaxID=9521;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tatesumi M., Hashimoto O.;
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
 CC receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck (by similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D86588; BAA131.1; -
 DR HSPB; P01730; 1M7O.
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.
 DR GO; GO:0015026; F:coreceptor activity; ISS.
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.
 DR GO; GO:0006955; P:immune response; ISS.
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
 DR GO; GO:0030217; P:T-cell differentiation; ISS.
 DR GO; GO:0045058; P:T-cell selection; ISS.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
 DR InterPro: IPR000973; CD4 TCAG.
 DR InterPro: IPR007110; IG-like.
 DR InterPro: IPR003596; IG_v.
 DR Pfam; PF00047; IG; 2.
 DR PRINTS; PR00692; CD4TCANTIGEN.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
 KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
 KW SIGNAL 1 25
 FT CHAIN 26 457
 FT DOMAIN 26 395
 FT TRANSMM 396 417
 POTENTIAL.

```

FT DOMAIN 418 457 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 26 125 IG-LIKE V-TYPE.
FT DOMAIN 126 202 IG-LIKE C2-TYPE 1.
FT DOMAIN 203 316 IG-LIKE C2-TYPE 2.
FT DOMAIN 317 373 IG-LIKE C2-TYPE 3.
FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 41 109 BY SIMILARITY.
FT DISULFID 155 184 BY SIMILARITY.
FT DISULFID 327 369 BY SIMILARITY.
FT LIPID 418 418 S-palmitoyl cysteine (By similarity).
FT LIPID 421 421 S-palmitoyl cysteine (By similarity).
SQ SEQUENCE 457 AA; 50871 MW; 57BED634405A015 CRC64;

Query Match 68.1%; Score 1582.5; DB 1; Length 457;
Best Local Similarity 69.9%; Pred. No. 3.8e-101;
Matches 320; Conservative 34; Mismatches 65; Indels 39; Gaps 4;

QY 1 MNRGVPFRHLIVLQALPPAATGKNVVLGKGGDTVLTCTASQKSIQFHMKNNOIK 60
DB 1 MNGGIPFRHLIVLQALPPAATGKTVLGGKGVVELPCTSLKKNVPRHMKTSQIK 60
QY 61 ILNGSGSFLTKGPSKLNDRADSRSLMDQGNFPLITKILKTEDSDTYICEVEDQKEEYQL 120
DB 61 ILGVNVTYTRGQSLTDRIDSKSSWDGSPFLIKARIEDSTYICEVESKKEEVEL 120
QY 121 LVFGLTANSDTHLLQGSQSLTLTLESPPSSPVQCSRSPRGKNIQGGKTLVSQLELQDSG 180
DB 121 QVFGLTANPDTHLLQGSQSLTLTLESPPSSPEVECTSPRGKIRIRKRTLSVQLCIPDSG 180
QY 181 TMTCTVLONOKKVEKIDIVLAPOKASSIYKKEGEVSEFPLAFVEKLTGSGELMW 240
DB 181 TWKCTVFOHLELV-EEINIVLAFOQASSTYKKEGEVSEFPLAFVEKLTGSGELCW 239
QY 241 QAERASSSSKSWITFDLKNKEVSVKRVTDOPKLQMGKKLPLHLTLFQALPOYAGSNTLTA 300
DB 240 QAERASSSSKSWITFDLQGEVYKLVTDOPKLQMGKKLPLHLTLFQALPOYAGSNTLTA 299
QY 301 LEAKTGKLGHOENVLVVNRATQLOKNLTCGEVWGFTSPKMLSLKENKEAKVSKREKPVW 360
DB 300 LKGGTGKLGHOENVLVVNRATQLOKNLTCGEVWGFTSPKMLSLKENQEAQVSKREKAVW 359
QY 361 LNPEAGMOCQLSDSGOVLLESNINIVLPTWSPVPPRPSALPAPTGSALDPOYASALP 420
DB 360 LNPEAGMOCQLSDSGOVLLESKFALPTRSPVQ----- 394
QY 421 DPAPASALPALAVISPLGL----GLGV-ACVLARTR 453
DB 395 -----PMVLIVGVAGLAFGLGIFLCVRCRHR 424

RESULT 11
CD4_RABIT STANDARD; PRT; 459 AA.
AC P46630.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE T4/Leu-3).
GN CD4.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxId=9986;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=92390370; PubMed=1518821;
RA Hage B.F., Sawadikiosol S., Brown T.J., Lee K., Recker D.P.,
RA Kindt T.J.,
RT "CD4 and its role in infection of rabbit cell lines by human
RT immunodeficiency virus type 1.";
```

```

RL Proc. Natl. Acad. Sci. U.S.A. 89:7963-7967(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
```

```

DR EMBL; M92840; AAA31198.1; -.
DR PIR; A46254; A46254.
DR HSSP; P01730; 1WBR.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro; IPR000973; CD4_TcAg.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TcANTIGEN.
DR SMART; SM00406; IGV_1.
DR PROSITE; PS50835; Ig_Like; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 459
FT DOMAIN 26 396
FT TRANSMEM 397 419
FT DOMAIN 420 459
FT DOMAIN 26 129
FT DOMAIN 130 208
FT DOMAIN 209 318
FT DOMAIN 319 374
FT CARBOHYD 299 299
FT DISULFID 41 113
FT DISULFID 329 370
FT LIPID 420 420
FT LIPID 423 423
SQ SEQUENCE 459 AA; 50886 MW; B523311C6D40013D CRC64;

Query Match 49.3%; Score 1147; DB 1; Length 459;
Best Local Similarity 58.0%; Pred. No. 2.5e-71;
Matches 240; Conservative 68; Mismatches 94; Indels 12; Gaps 6;

QY 1 MNRGVPFRHLIVLQALPPAATGKNVVLGKGGDTVLTCTASQKSIQFHMKNNOIK 60
DB 1 MNRGVPFRHLIVLQALPPAATGKNVVLGKGGDTVLTCTASQKSIQFHMKNNOIK 60
QY 61 ILNGSGSFLTKGPSKLNDRADSRSLMDQGNFPLITKILKTEDSDTYICEVEDQKEEYQL 116
DB 61 ILNGSGSFLTKGPSKLNDRADSRSLMDQGNFPLITKILKTEDSDTYICEVEDQKEEYQL 116
QY 117 EVOLVFGTANSDTHLLQGSQSLTLTLESPPSSPVQCSRSPRGKNIQGGKTLVSQLEL 176
DB 117 EVOLVFGTANSDTHLLQGSQSLTLTLESPPSSPVQCSRSPRGKNIQGGKTLVSQLEL 176
QY 121 EVELLVFRLTANPTRLHLHGQSLTLTLEGPVSGPSVQWSPKRIETGPTGMPRLRL 180
DB 121 EVELLVFRLTANPTRLHLHGQSLTLTLEGPVSGPSVQWSPKRIETGPTGMPRLRL 180
QY 177 QDSGTWCTVLTQVQKVEKIDIVLAFOKASSIYKKEGEVSEFPLAFVEKLTGSG 235
DB 177 QDSGTWCTVLTQVQKVEKIDIVLAFOKASSIYKKEGEVSEFPLAFVEKLTGSG 235
QY 181 QDSGTWCTVLTQVQKVEKIDIVLAFOKASSIYKKEGEVSEFPLAFVEKLTGSG 238
DB 181 QDSGTWCTVLTQVQKVEKIDIVLAFOKASSIYKKEGEVSEFPLAFVEKLTGSG 238
QY 236 GELMWQAERASSSSKSWITFDLKNKEVSVKRVTDOPKLQMGKKLPLHLTLFQALPOYAGSG 295
```

```

DB 239 GELMWOVQAGASSAGSWFSFLEDRKVSQKILPDLKIQMSKGLPLSLTLPOLARHYAGSG 298
OY 296 NLTLLALEAKTKLHOENVLVNMRATQLOKNTLCEVWGTSFKMLSLKLENKEAKVSKRE 355
DB 299 NLSLTLD--KGLHQVSLVMLKVTVQKNKLTCEVLGIDPKMLSLKLEDEQEAQVS-TQ 355
OY 356 KPVWLVNDEAGMOCCLSDSGVLLSESNIKVLPWSTFVHPRASALPAPPTGSA 409
DB 356 KMOVLDKPKAGTKWQCLSSGQVLLSEKADVLATGLS--HQOPTLLAGALGTA 407

RESULT 12
ID CD4_CANFA STANDARD; PRT; 463 AA.
AC P33705;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN CD4
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE OF 13-463 FROM N.A.
RC STRAIN=Beagle; TISSUE=Thymus;
RX MEDLINE=93192324; PubMed=7916632;
RA Milde K.F., Conner G.E., Minz D.H., Alejandro R.;
RT "Primary structure of the canine CD4 antigen.";
RL Biochim. Biophys. Acta 1172:315-318(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Thymus;
RX MEDLINE=94378217; PubMed=8091416;
RA Gorman S.D., Frewin M.R., Cobbold S.P., Waldmann H.;
RT "Isolation and expression of cDNA encoding the canine CD4 and CD8
RT alpha antigens."
RL Tissue Antigens 43:184-188(1994).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in macrophages and a subset of
CC T lymphocytes.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC
DB EMBL; L06130; AB02295.1; -
DB EMBL; X68565; -, NOT_ANNOTATED_CDS.
DB HSSP; P01730; 1MR.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR00973; CD4_TGAG.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 3.

```

```

DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IG; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1
FT CHAIN 25
FT DOMAIN 25 463
FT TRANSMEM 402 423
FT DOMAIN 424 463
FT DOMAIN 26 124
FT DOMAIN 125 211
FT DOMAIN 212 321
FT DOMAIN 322 378
FT DISULFID 41 109
FT DISULFID 332 374
FT LIPID 424
FT LIPID 427 424
FT CARBOHYD 123 123
FT CARBOHYD 168 168
FT CARBOHYD 176 176
FT CARBOHYD 324 324
FT CARBOHYD 329 329
FT CARBOHYD 389 389
SQ SEQUENCE 463 AA; 51639 MW; 95805170CB4A833 CRC64;

Query Match 48.94; Score 1137; DB 1; Length 463;
Best Local Similarity 52.54; Pred. No. 1.2e-70;
Matches 242; Conservative 67; Mismatches 100; Indels 52; Gaps 8;

OY 1 MNRGVPFRLHLVLIQALPAPATQGNKVVVGKKGDTVELTCTASQKSLQPHMKNNOIK 60
DB 1 MNRGVPFRLHLVLIQALPAPATQGNKVVVGKKGDTVELTCTASQKSLQPHMKNNOIK 60
OY 61 ILNQGSEFLTKGSKLNDRAKRSRLMDQNFLLIKNTKLESDSYTYICEVDEQKEVQL 120
DB 61 ILNQGSEFLTKGSKLNDRAKRSRLMDQNFLLIKNTKLESDSYTYICEVDEQKEVQL 120
OY 61 ILNQGSEFLTKGSKLNDRAKRSRLMDQNFLLIKNTKLESDSYTYICEVDEQKEVQL 120
DB 61 ILNQGSEFLTKGSKLNDRAKRSRLMDQNFLLIKNTKLESDSYTYICEVDEQKEVQL 120
OY 121 LVFGLT-----NSDTHLQGGSLTLESPGSSPSVOCRSPRGNIOGKTLV 171
DB 121 LVFGLT-----NSDTHLQGGSLTLESPGSSPSVOCRSPRGNIOGKTLV 171
OY 120 LVFNLTKMDSSGSSSNRLQGGQLTTLNPSGSSPSVWKGPKGSKGGQNL 179
DB 120 LVFNLTKMDSSGSSSNRLQGGQLTTLNPSGSSPSVWKGPKGSKGGQNL 179
OY 172 SOLELDPSGTWTCTVLQNKKEVFKIDIVLAFOKASIVYKKEGQVSEFPLAFTVEK 231
DB 172 SOLELDPSGTWTCTVLQNKKEVFKIDIVLAFOKASIVYKKEGQVSEFPLAFTVEK 231
OY 180 SWPELDQGTWTCITISGQKTEFVNVLAVQKSNFTYAREGQDVERSFPLSFEDEN 239
DB 180 SWPELDQGTWTCITISGQKTEFVNVLAVQKSNFTYAREGQDVERSFPLSFEDEN 239
OY 232 LTSGEILMWAQEAASSKSWITFDLKNKEVSVKRVTPDKLQNGKPLHLTLPOLPOY 291
DB 232 LTSGEILMWAQEAASSKSWITFDLKNKEVSVKRVTPDKLQNGKPLHLTLPOLPOY 291
OY 240 LV--GELRWQAQAGASSSLMISEFTLENRKLMSKEAHAPLKLQMKESLPRLTPOVL 297
DB 240 LV--GELRWQAQAGASSSLMISEFTLENRKLMSKEAHAPLKLQMKESLPRLTPOVL 297
OY 292 AGSGNLTLLALEAKTKLHOENVLVNMRATQLOKNTLCEVWGTSFKMLSLKLENKEAKV 351
DB 292 AGSGNLTLLALEAKTKLHOENVLVNMRATQLOKNTLCEVWGTSFKMLSLKLENKEAKV 351
OY 298 AGSGILTLNI-AK-GTLYQENVLVNMRANSSQNNLTCEVGLPSTPILTLNLKEQAKV 355
DB 298 AGSGILTLNI-AK-GTLYQENVLVNMRANSSQNNLTCEVGLPSTPILTLNLKEQAKV 355
OY 352 SKKEKPVWLVNDEAGMOCCLSDSGVLLSESNIKVLPWSTFVHPRASALPAPPTGSA 411
DB 352 SKKEKPVWLVNDEAGMOCCLSDSGVLLSESNIKVLPWSTFVHPRASALPAPPTGSA 411
OY 356 SKQKQLVWVDPGEGTWQCLSDKXVLLASSLNV-----SSPV----- 394
DB 356 SKQKQLVWVDPGEGTWQCLSDKXVLLASSLNV-----SSPV----- 394
OY 412 DPQTASALPDPAPASALPALA-----VISFLGLGLGVNC 447
DB 412 DPQTASALPDPAPASALPALA-----VISFLGLGLGVNC 447
OY 395 -----VIKSWPKFLATITGLIGLLIGLCVFC 423
DB 395 -----VIKSWPKFLATITGLIGLLIGLCVFC 423

RESULT 13
ID CD4_RAT STANDARD; PRT; 457 AA.
AC P05540;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3) (W3/25 antigen).
GN CD4.

```

OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxId=10116;
 RN
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=87175535; PubMed=3104900;
 RA Clark S.J., Jeffries W.A., Barclay A.N., Gagnon J., Williams A.F.,
 RT "Peptide and nucleotide sequences of rat CD4 (W3/25) antigen:
 RT evidence for derivation from a structure with four
 RT immunoglobulin-related domains.";
 RT Proc. Natl. Acad. Sci. U.S.A. 84:1649-1653(1987).
 RN
 RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 210-399.
 RX MEDLINE=93262437; PubMed=8493555.
 RA Brady R.L., Dodson E.J., Dodson G.G., Lange G., Davis S.J.,
 RA Williams A.F., Barclay A.N.;
 RT "Crystal structure of domains 3 and 4 of rat CD4: relation to the
 RT NH2-terminal domain.";
 RT Science 260:979-983(1993).
 CC
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
 CC receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M15768; AAA0901.1; -
 DR PIR; A27449; A27449.
 DR PDB; 1CID; 15-JUL-93.
 DR
 DR GLYCOSITE: P05540.
 DR GO; GO:0042101; C1T-cell receptor complex; ISS.
 DR GO; GO:0015026; F:coreceptor activity; ISS.
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.
 DR GO; GO:0006955; P:immune response; ISS.
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
 DR GO; GO:0030217; P:T-cell differentiation; ISS.
 DR GO; GO:0045058; P:T-cell selection; ISS.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
 DR InterPro: IPR000973; CD4 TCAG.
 DR InterPro: IPR007110; Ig-Like.
 DR InterPro: IPR003599; Ig.
 DR Pfam; PF00047; Ig_2.
 DR PRINTS; PR00692; CD4TCANTIGEN.
 DR SMART; SM00409; IG_2.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR
 DR Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
 KW Immune response; Repeat; Signal; Lipoprotein; Palmitate; 3D-structure.
 FT SIGNAL 1 27
 FT CHAIN 28 457
 FT DOMAIN 28 394
 FT TRANSMEM 395 417
 FT DOMAIN 418 457
 FT DOMAIN 28 127
 FT DOMAIN 128 206
 FT DOMAIN 207 316
 FT DOMAIN 317 374
 FT CARBOHYD 186 186
 FT CARBOHYD 297 297
 FT CARBOHYD 392 392
 FT DISULFID 43 111
 FT DISULFID 158 187
 FT DISULFID 370 370
 FT LIPID 418 421
 FT LIPID 421 421

FT STRAND 213 217
 FT TURN 218 219
 FT STRAND 222 225
 FT STRAND 235 243
 FT STRAND 252 258
 FT TURN 259 260
 FT STRAND 261 265
 FT STRAND 274 275
 FT STRAND 278 280
 FT TURN 279 280
 FT STRAND 282 285
 FT HELIX 290 292
 FT STRAND 294 301
 FT STRAND 306 319
 FT STRAND 325 331
 FT STRAND 338 344
 FT TURN 345 346
 FT STRAND 349 353
 FT STRAND 357 361
 FT STRAND 367 374
 FT TURN 375 376
 FT STRAND 377 385
 FT SEQUENCE 457 AA; 51437 MW; 477BE157D30954C1 CRC64;
 SO
 Query Match 43.0%; Score 999; DB 1; Length 457;
 Best Local Similarity 52.9%; Pred. No. 3.4e-61;
 Matches 207; Conservative 61; Mismatches 115; Indels 8; Gaps 5;
 QY 1 MNRGVPRHL--LLVLQALLPAATQGNKVVLGKGDVLELTCTASQKKSIOFHWKSNQ 58
 DB 1 MCRGFSRHLPLPILLOSLKLVLTQKTVVLGKGGSAELPCSTSRSSAFMKSSDQ 60
 QY 59 IKIIGNQSFTKPKSKLNDPDRSRSLMDQGNPILIKLIKIDSDTYITCEVDQKEV 118
 DB 61 KTIIGYKXKLLIKSLSELYSRFDRKNAWEGSPILLINKRMEDSOTYVCELNKKEV 120
 QY 119 QLVFGLTANSDTYHLOQGSITLTLES-PRGSSPVOCRSPRGKNIQGGKTLVSQLEQ 177
 DB 121 ELWFRVTFNPGTILLOQGSITLTLDNSPKVSDPIECRKHSSIVVDSKAFSTHSLRIQ 180
 QY 178 DSGTWTCTVLONQKKEFKIDIVLAFQKASSIYKKEGEQVEFSFPLAFTVEKLTSGGE 237
 DB 181 DSGIINCTVLTNQKHSFDMKLSVLGFASSTITAYKSEGEAEFSFPLNIGEESL--QGE 238
 QY 238 LMQAERASSKSSKITTPELDRKKEVSVKRVTDPELQMGKPLHLTPALPQVAGSNTL 297
 DB 239 LRMFAEKAPSSQSWITPSLKNQKVSVOKSTSNPKFQLESLPLTLQIPOVSLQFAGSGNTL 298
 QY 298 TLAEAKTKLHOBVNLVVMRATOLQKN-LTCEVWGPSPPLMLSLKLENKEAKVSKREK 356
 DB 299 TLTIID--RGILYQENVNLVVMKVTOPDSNTLTCEVWGPSPPLMLSLKLENKEAKVSKREK 356
 QY 357 PFWVLNPEAGMWOGLSDSGQVLLSNIKYL 387
 DB 357 VIQVAPBAGVWQGLLSGGEVKNMDSKIQVL 387
 RESULT 14
 CD4_MOUSE
 ID CD4_MOUSE STANDARD; PRT; 457 AA.
 AC P06332;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
 DE T4/Leu-3) (T-cell differentiation antigen L3T4).
 GN CD4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxId=10090;
 RN
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=87018645; PubMed=3094146;
 RA "Touvalille B., Gorman S.D., Field E.H., Hunkapiller T., Parnes J.R.;
 RT "Isolation and sequence of L3T4 complementary DNA clones: expression
 RL in T cells and brain.";
 RN Science 234:610-614(1986).
 [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=87115821; PubMed=3027575;
 RA Littman D.R., Gettner S.N.;
 RT "Unusual intron in the immunoglobulin domain of the newly isolated
 RL murine CD4 (L3T4) gene.";
 RN Nature 325:453-455(1987).
 [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=88152875; PubMed=3326818;
 RA Parnes J.R., Hunkapiller T.;
 RT "L3T4 and the immunoglobulin gene superfamily: new relationships
 RL between the immune system and the nervous system.";
 RN Immunol. Rev. 100:109-127(1987).
 [4]
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RX TISSUE=Brain;
 RA MEDLINE=88041159; PubMed=2823269;
 RA Gorman S.D., Touvalille B., Parnes J.R.;
 RT "Structure of the mouse gene encoding CD4 and an unusual transcript
 RL in brain.";
 RN Proc. Natl. Acad. Sci. U.S.A. 84:7644-7648(1987).
 [5]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=98112780; PubMed=9445485;
 RA Anselmi-Lati M.A., Oeltjen J.C., Schwartz S., Zhang Z., Muzny D.M.,
 RA Lu J., Gorrell J.H., Chnault A.C., Belmont J.W., Miller W.,
 RA Gibbs R.A.;
 RT "Comparative sequence analysis of a gene-rich cluster at human
 RL chromosome 12p13 and its syntenic region in mouse chromosome 6.";
 RN Genome Res. 8:29-40(1998).
 [6]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RX STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RA MEDLINE=22388257; PubMed=12477932;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stepleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein W.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kertman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Buterfield A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences.";
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [7]
 RN SEQUENCE OF 27-43.
 RX MEDLINE=8616694; PubMed=3082751;
 RA Classon B.J., Tsagaratos J., Kirschbaum L., Maddox J., McKay C.R.,
 RA Brandon M., McKenzie I.F.C., Walker I.D.;
 RT "The L3T4 antigen in mouse and the sheep equivalent are
 RL immunoglobulin-like.";
 RN Immunogenetics 23:129-132(1986).
 [8]
 RX DISULFIDE BONDS.
 RA MEDLINE=86233454; PubMed=3086886;
 RA Classon B.J., Tsagaratos J., McKenzie I.F.C., Walker I.D.;
 RT "The L3T4 antigen in mouse and the sheep equivalent are
 RL immunoglobulin-like.";
 RN Immunogenetics 23:129-132(1986).

RT "Partial primary structure of the T4 antigens of mouse and sheep:
 RL assignment of intrachain disulfide bonds.";
 RN Proc. Natl. Acad. Sci. U.S.A. 83:4499-4503(1986).
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
 CC receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck (by similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=P06332-1; Sequence=displayed;
 CC Name=2; Synonyms=Brain-specific;
 CC IsoId=P06332-2; Sequence=VSP_002489;
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M36850; AAA39401.1; -;
 CC EMBL; M13816; AAA37267.1; -;
 CC EMBL; X04836; CAA28539.1; -;
 CC EMBL; M36851; AAA39402.1; -;
 CC EMBL; M17080; AAA37403.1; -;
 CC EMBL; M17078; AAA37403.1; JOINED.
 CC EMBL; M17079; AAA37403.1; JOINED.
 CC EMBL; BC039397; AAC36010.1; -;
 CC EMBL; BC039397; AAC39337.1; -;
 CC PIR; A02110; RWMST4.
 CC HSP; P01730; IMR.
 CC MGD; MGI:88335; Cd4.
 CC GO; GO:0042101; C:T-cell receptor complex; ISS.
 CC GO; GO:0015026; F:coreceptor activity; ISS.
 CC GO; GO:0042289; F:MHC class II protein binding; ISS.
 CC GO; GO:0006955; P:immune response; ISS.
 CC GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
 CC GO; GO:0030217; P:T-cell differentiation; ISS.
 CC GO; GO:0045058; P:T-cell selection; ISS.
 CC GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
 CC InterPro; IPR000973; CD4_TcRg.
 CC InterPro; IPR007110; Ig-Like.
 CC InterPro; IPR003596; Ig_V.
 CC Pfam; PF00047; Ig_2.
 CC PRINTS; PR00692; CD4TCANTIGEN.
 CC SMART; SM00406; IGV_1.
 CC PROSITE; PS50835; IG_LIKE_1.
 CC Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
 CC Immune response; Repeat; Signal; Lipoprotein; Palmitate;
 CC Alternative splicing.
 CC KW
 CC CHAIN 1
 CC DOMAIN 27 457
 CC TRANSMEM 27 394
 CC DOMAIN 395 417
 CC DOMAIN 418 457
 CC DOMAIN 27 128
 CC DOMAIN 129 207
 CC DOMAIN 208 317
 CC DOMAIN 318 374
 CC CARBOHYD 187 187
 CC CARBOHYD 298 298
 CC CARBOHYD 323 323
 CC CARBOHYD 392 392
 CC DISULFID 42 112
 CC DISULFID 159 188
 CC DISULFID 328 370
 CC LIPID 418 418
 CC LIPID 421 421
 CC VARSPIC 1 240
 CC
 CC T-CELL SURFACE GLYCOPROTEIN CD4.
 CC EXTRACELLULAR (POTENTIAL).
 CC POTENTIAL.
 CC CYTOPLASMIC (POTENTIAL).
 CC IG-LIKE V-TYPE.
 CC IG-LIKE C2-TYPE 1.
 CC IG-LIKE C2-TYPE 2.
 CC IG-LIKE C2-TYPE 3.
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC S-palmitoyl cysteine (By similarity).
 CC S-palmitoyl cysteine (By similarity).
 CC Missing (in isoform 2).

```

FT SEQUENCE 457 AA; 51296 MW; 181DA7527CB00F33 CRC64; /FTId=VSP_002489.
SQ Query Match 42.7%; Score 993; DB 1; Length 457;
Beet Local Similarity 53.8%; Pred. No. 8,8e-61;
Matches 211; Conservative 62; Mismatches 109; Indels 10; Gaps 7;

QY 1 MNRGVPFRF-LTLVQLALPPAATGKNKVLKGGDPTVLTCTAQQKSIQPHMKNQI 59
DB 1 MCRALSLRLHLHLQLQLAVTQKTLVLGSGSABLPCESSQKITYFTWKFSDR 60
QY 60 KILGNGQ-SFLTKG--PSKLRADRSRSLMDQGNFPLIKNLKIEDSPYICEVEDQKE 116
DB 61 KILGNGKGVLRGSGPSQF-DRPDSKKGAMKSGPPLINKLKNMEDSTYICELNNKE 119
QY 117 EVQLVFLGILTANSDFHLLOGSLTLTLES-PPGSSPSVQCSPPKKNIOGKTLVSQLE 175
DB 120 EVELWVFVKTFSPGTSLLQGSGLTLTLDNSKVSNPDLTECKRKKGVSGSKVLSMSNLR 179
QY 176 LQDSGWTCTVLQONKKVEFKIDIVLAFQKASITVYKKGQVESPFLAFTVEKLGS 235
DB 180 VQDSDFMNCVTVLQDKRMFGMTLSVLFQSTAIRAYSGESABSPFLNFAER--NGW 237
QY 236 GELMWQAERASSSKSWITFDLKNKEYSVRVTQDPRLQWKKLPLHLTLPLQALPYAGSG 295
DB 238 GELMWKAEDSFQFPQIPISFISINKKSVQSKTKDLQKLTPLPLIKIPQVSLQFASG 297
QY 236 NITLLEAKTGKLVHQRNLVVRAPQLQKNTLCEVWGPSPKLMLSLKLENKAVSKRE 355
DB 238 NITLTLTD--KGTILHDEVNLVWVKVAQLNNTLTCEVWGPSPKMRILTLKQENQEARVSEBQ 355
QY 356 KPVWLVLPKAGMWQCLSDSGQVLLSENIKVL 387
DB 356 KVVQVAPETGLMQCLSEGDKVKNDSRIQVL 387

RESULT 15
CD7_HUMAN STANDARD; PRT; 240 AA.
AC P09564;
DT 01-MAR-1989 (rel. 10, Created)
DT 01-MAR-1989 (rel. 10, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE T-cell antigen CD7 precursor (GP40) (T-cell leukemia antigen) (TP41)
DE (Leu-9).
GN CD7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN 1[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8811517; PubMed=3501369;
RA Aruffo A., Seed B.;
RT "Molecular cloning of two CD7 (T-cell leukemia antigen) cDNAs by a
RT COS cell expression system.";
RL EMBO J. 6:3313-3316(1987).
RN 2[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9110576; PubMed=1703303;
RA Schanberg L.E., Fleener D.E., Kurtzberg J., Haynes B.F., Kaufman R.E.;
RT "Isolation and characterization of the genomic human CD7 gene:
RT structural similarity with the murine Thy-1 gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:603-607(1991).
RN 3[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schenfer C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buecker K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

```

```

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.A., Toshiyuki S., Carrini P., Prange C.,
RA Raha S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gamarate P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Characterization and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN 4[4]
RP SEQUENCE OF 205-240 FROM N.A.
RX MEDLINE=91267564; PubMed=1711009;
RA Yoshikawa K., Seto M., Ueda R., Obata Y., Notake K., Yokochi T.,
RA Takahashi T.;
RT "Molecular cloning of the gene coding for the human T cell
RT differentiation antigen CD7.";
RL Immunogenetics 33:352-360(1991).
RN 5[5]
RP TOPOLOGY.
RX MEDLINE=90063052; PubMed=2479685;
RA Ware R.E., Searce R.M., Dietz M.A., Starmer C.F., Parker T.J.,
RA Haynes B.F.;
RT "Characterization of the surface topography and putative tertiary
RT structure of the human CD7 molecule.";
RL J. Immunol. 143:3632-3640(1989).
RN 6[6]
RP INTERACTION WITH SECTM1.
RX MEDLINE=20119303; PubMed=10652336;
RA Lyman S.D., Escobar S., Rousseau A.-M., Armstrong A., Fanslow W.C.;
RT "Identification of CD7 as a cognate of the human K12 (SECTM1)
RT protein.";
RL J. Biol. Chem. 275:3431-3437(2000).
RN 7[7]
RP FUNCTION: Not yet known.
RN 8[8]
RP SUBUNIT: Interacts with SECTM1.
RN 9[9]
RP SIMILARITY: Contains 1 immunoglobulin-like domain.
RN 10[10]
RP DATABASE: NAME=PROW; NOTE=CD guide CD7 entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd7.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; X06180; CAA29546.1; -
DR EMBL; BC37271; AAA51953.1; -
DR EMBL; M3009293; AAA09293.1; -
DR EMBL; BC013297; AAA13297.1; -
DR EMBL; D00749; BAA00646.1; -
DR EMBL; D00747; BAA00646.1; JOINED.
DR EMBL; D00748; BAA00646.1; JOINED.
DR PIR; A39016; A39016.
DR HSRP; P01607; IREI.
DR Genew; HGNC:1695; CD7.
DR MIM; 186820; -
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005624; C:membrane fraction; TAS.
DR GO; GO:0005866; C:plasma membrane; TAS.
DR GO; GO:0004872; F:receptor activity; TAS.
DR GO; GO:0006955; P:immune response; TAS.
DR GO; GO:0042110; P:T-cell activation; TAS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. ; NNS.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003599; IG.

```

```
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW T-cell; Signal; Immune response; Antigen; Transmembrane; Glycoprotein;
KW Immunoglobulin domain; Receptor; Lipoprotein; Palmitate; Repeat.
FT SIGNAL 1 25
FT CHAIN 26 240
FT DOMAIN 26 180 T-CELL ANTIGEN CD7.
FT TRANSMEM 181 201 EXTRACELLULAR (PROBABLE).
FT DOMAIN 202 240 PROBABLE.
FT DOMAIN 26 130 CYTOPLASMIC (PROBABLE).
FT DOMAIN 145 180 IG-LIKE.
FT 4 x 9 AA TANDEM REPEATS, POTENTIAL SPACER
FT FUNCTION.
FT DISULFID 35 142
FT DISULFID 48 114 POTENTIAL.
FT LIPID 198 198 S-palmitoyl cysteine.
FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .).
FT REPEAT 145 153 1.
FT REPEAT 154 162 2.
FT REPEAT 163 171 3.
FT REPEAT 172 180 4.
SQ SEQUENCE 240 AA; 25409 MW; EBBCE08279552108 CRC64;

Query Match 12.6%; Score 293; DB 1; Length 240;
Best Local Similarity 40.3%; Pred. No. 3.3e-13;
Matches 91; Conservative 18; Mismatches 49; Indels 68; Gaps 9;

QY 270 PRLQWKKLPHLTPALPOVAGSGLTLALAKTGKHOENVLYVMRATQLOKNLTCE 329
DB 5 PRLTL---LPLLLALARGLPG-----ALAAQEVQOSPCHTTPVGAS---VNITCS 49
QY 330 VWGPTSPKLM.LSLKLENKAVSKREKPVWVNLNPEAG----- 366
DB 50 TSGGARGVYLNQL-----GPQDIIYVDGVPTTDRFRGRIDFSGSQDNL 98
QY 367 --MMOCLISDSG---QVLESNT-----KVLPT-----N-STVHPRASALPAPPTG 407
DB 99 ITMRLQLSLDTGTVCQAITEVNVSGTLVLYVEEQQGMRHRCSDADPPRASALPAPPTG 158
QY 408 SALPDPOTASALPDPASALPALAVISFLIGLGVACVLTARR 453
DB 159 SALPDPOTASALPDPASALPALAVISFLIGLGVACVLTARTQ 204

RESULT 16
VCAL_RAT ID _VCAL_RAT STANDARD; PRT; 739 AA.
AC P29534;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Vascular cell adhesion protein 1 precursor (V-CAM 1).
OS NCAM1 OR VCAM-1.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RC MEDLINE=92181437; PubMed=1371918;
RA Hession C., Moy P., Tizard R., Chisholm P., Williams C., Wyek M.,
RA Burkle L., Miyake K., Kincaid P., Lobb R.;
RT "Cloning of murine and rat vascular cell adhesion molecule-1.";
RL Biochem Biophys Res Commun 183:163-169(1992).
CC -1- FUNCTION: IMPORTANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION
CC IN LEUKOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1
CC INTEGRIN VLA4 ON LEUKOCYTES, AND MEDIATES BOTH ADHESION AND SIGNAL
CC TRANSDUCTION. THE VCAM1/VLA4 INTERACTION MAY PLAY A
CC PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUKOCYTE
CC EMIGRATION TO SITES OF INFLAMMATION.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
```

```
CC -1- TISSUE SPECIFICITY: Expressed on inflamed vascular endothelium, as
CC well as on macrophage-like and dendritic cell types in both normal
CC and inflamed tissue.
CC -1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; M84488; AAA42332.1; -.
CC PIR; J50675; J50675.
CC HSSP; P19320; IYCA.
CC InterPro; IPR003987; ICAM VCAM-1.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR003989; VCAM-1.
CC Pfam; PF00047; Ig; 5.
CC PRINTS; PR01472; ICAMVCAM1.
CC PRINTS; PR01474; VCAM1.
CC SMART; SM00408; IGc2; 3.
CC PROSITE; PS50835; IG_LIKE; 5.
KW Immunoglobulin domain; Glycoprotein; Cell adhesion; Transmembrane;
KW Repeat; Signal.
FT SIGNAL 1 24 PROBABLE.
FT CHAIN 25 739 VASCULAR CELL ADHESION PROTEIN 1.
FT DOMAIN 25 698 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 699 720 POTENTIAL.
FT DOMAIN 721 739 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 119 212 IG-LIKE C2-TYPE 1.
FT DOMAIN 223 309 IG-LIKE C2-TYPE 2.
FT DOMAIN 312 397 IG-LIKE C2-TYPE 3.
FT DOMAIN 408 506 IG-LIKE C2-TYPE 4.
FT DOMAIN 514 595 IG-LIKE C2-TYPE 5.
FT DOMAIN 601 682 IG-LIKE C2-TYPE 6.
FT DISULFID 47 95 BY SIMILARITY.
FT DISULFID 52 99 BY SIMILARITY.
FT DISULFID 137 195 BY SIMILARITY.
FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 424 424 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 561 561 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 650 650 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 739 AA; 81246 MW; 5C6085A1A1B100C CRC64;

Query Match 7.0%; Score 162; DB 1; Length 739;
Best Local Similarity 22.0%; Pred. No. 0.0013;
Matches 86; Conservative 60; Mismatches 155; Indels 90; Gaps 14;

QY 33 KEDVELITCTASQKKSQIFWKNKSNQIKLGNQSFITKPSGLNDRADRSRLMDGNF 92
DB 238 EGAAVTTCASSEGUPAPEIFWSKK-----LDNGVLQDL-----SGNA 274
QY 93 PLIKNIKIDSDTYICE-----VEDQKEVQLV-----FGLTANSDFTLQSGSLTLT 142
DB 275 TLTILAKMEDSGIYCEGNLVGRDKTEVELLVQEKRFVVDISPSQVLAQVGDVVL 334
QY 143 LESDPGSPSVQCRSPKKNIGQ-----GKTLVSQLELDQSGTWTCTVLQNKQKVEF 195
DB 335 CAAVAGCSPSPFMSRQTQDSPLNGBVNDGATSTLTLSPVGEVBHSLCTVTCQRRLEK 394
QY 196 KIDIVLAFQKASIVYKKGEOVSPSPPLA-----FTVEKLTGSG 236
DB 395 TIQEVYSF-----PDPEIETISGPLVHGRPVTVNCTVNVVPFDHLEIELLNGET 445
QY 237 ELWVQAEARRAS-SKSWITFLPKNKEVSVKRVTDPPKLGWKKL-----PLHLTPQALP-Q 290
DB 446 TLANKFLREITGKTS---LETKSLNMTFTPAED---TGKALVCLAKLHSSQWSESPKQ 498
```

```

OY 291 VAGSGLTALBPAKTKGLHQEVNLVYMRATQKQKLTCEVWPTSPKMLSTKLENKAK 350
DB 499 ROSTQTLVYVNAKPKPTIVWSPSPVPEESPP--NLTCSDGPFPPKILMSQKNGELQ 556
OY 351 VSKREKPVVNLPEAGMOCILSDSGQVILE 381
DB 557 PLSQ-----NTTSLFMATKEDSGIYVCE 580

RESULT 17
VCAL_HUMAN
ID _VCAL_HUMAN STANDARD; PRT; 739 AA.
AC P19320;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Vascular cell adhesion protein 1 precursor (V-CAM 1) (CD106 antigen)
DE (INCAM-100).
OS VCAM1 OR L1CAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=umbilical vein;
RX MEDLINE=91016951; PubMed=1699207;
RA Pole T., Newman W., Gopal T.V.;
RT "Full length vascular cell adhesion molecule 1 (VCAM-1).";
RL Nucleic Acids Res. 18:5901-5901(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90090619; PubMed=2668898;
RA Osborn L., Hesston C., Tizard R., Vassallo C., Luhnowsky S.,
RA Chi-Rosco G., Lobb R.;
RT "Direct expression cloning of vascular cell adhesion molecule 1, a
RT cytokine-induced endothelial protein that binds to lymphocytes.";
RL Cell 59:1203-1211(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=9152090; PubMed=1715583;
RA Cybulsky M.I., Fries J.W.U., Williams A.J., Sultan P., Eddy R.,
RA Byers M., Shows T., Gimbrone M.A. Jr., Collins T.;
RT "Gene structure, chromosomal location, and basis for alternative mRNA
RT splicing of the human VCAM1 gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:7859-7863(1991).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=91201302; PubMed=1707873;
RA Hesston C., Tizard R., Vassallo C., Schiffer S.B., Goff D., Moy P.,
RA Chi-Rosco G., Luhnowsky S., Lobb R., Osborn L.;
RT "Cloning of an alternate form of vascular cell adhesion molecule-1
RT (VCAM1).";
RL J. Biol. Chem. 266:6682-6685(1991).
RN [5]
RP SEQUENCE FROM N.A., AND VARIANTS PHE-318; ALA-384; ALA-413 AND
RP LEU-716.
RA Rieder M.J., Armet T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
RA Rajkumar N.R., Toch E.J., Yi Q., Nickerson D.A.;
RT Submitted (Aug-2002) to the EMBL/Genbank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC TISSUE=Retinal pigment epithelium;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abrahamsen R.D., Mullighy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

```

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalloe D.E.,
RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 25-226.
RX MEDLINE=95147978; PubMed=7531291;
RA Jones E.Y., Harlos K., Bottomley M.J., Robinson R.C., Driscoll P.C.,
RA Edwards R.W., Clements J.M., Dudgeon I.J., Stuart D.I.;
RT "Crystal structure of an integrin-binding fragment of vascular cell
RT adhesion molecule-1 at 1.8-A resolution.";
RL Nature 373:539-544(1995).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 25-220.
RX MEDLINE=95296382; PubMed=7539925;
RA Wang J.-H., Pepinsky R.B., Stehle T., Liu J.-H., Karpusas M.,
RA Browning B., Osborn L.;
RT "The crystal structure of an N-terminal two-domain fragment of
RT vascular cell adhesion molecule 1 (VCAM-1): a cyclic peptide based on
RT the domain 1 C-D loop can inhibit VCAM-1-alpha 4 integrin
RT interaction.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:5714-5718(1995).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 25-220.
RA Wang J.-H., Stehle T., Pepinsky R.B., Liu J.-H., Karpusas M.,
RA Osborn L.;
RT "Structure of a functional fragment of VCAM-1 refined at 1.9-A
RT resolution.";
RL Acta Crystallogr. D 52:369-379(1996).
CC -I- FUNCTION: IMPORTANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION
CC IN LEUKOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1
CC INTEGRIN VLA4 ON LEUKOCYTES, AND MEDIATES BOTH ADHESION AND SIGNAL
CC TRANSDUCTION. THE VCAM1/VLA4 INTERACTION MAY PLAY A
CC PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUKOCYTE
CC EMIGRATION TO SITES OF INFLAMMATION.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=Long;
CC IsoId=P19320-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P19320-2; Sequence=VSP_002580;
CC -I- TISSUE SPECIFICITY: Expressed on inflamed vascular endothelium, as
CC well as on macrophage-like and dendritic cell types in both normal
CC and inflamed tissue.
CC -I- INDUCTION: By cytokines (e.g. IL-1, TNF-alpha).
CC -I- PTM: Sialoglycoprotein.
CC -I- DISEASE: May play an important role in the genesis of
CC atherosclerosis and rheumatoid arthritis.
CC -I- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.
CC -I- DATABASE: NAME=PROV; NOTE=CD guide CD106 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd106.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X53051; CA37218.1; -
CC EMBL; M30257; AA51917.1; ALT TERM.
CC EMBL; M73355; AAA61270.1; -
CC EMBL; M60335; AAA61269.1; -

```

```

DR EMBL; AF536818; AAM96190.1; -.
DR EMBL; BC017276; AAH17276.1; -.
DR PIR; A41288; A41288.
DR PIR; B41288; B41288.
DR PDB; 1VCA; 20-JUN-96.
DR PDB; 1VSC; 15-SEP-95.
DR PDB; 1I09; 07-NOV-01.
DR Genew; HGNC:12663; VCAM1.
DR MIM; 192225; -.
DR InterPro; IPR003987; ICAM VCAM-1.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003588; Ig_C2.
DR InterPro; IPR003989; VCAM-1.
DR Pfam; PF00047; Ig; 6.
DR PRINTS; PR01472; ICAMVCAM1.
DR PRINTS; PR01474; VCAM1.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS50835; IG_LIKE; 5.
KW Immunoglobulin domain; Glycoprotein; Cell adhesion; Transmembrane;
Repeat; Signal; Alternative splicing; Polymorphism; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 1 25
FT DOMAIN 25 698
FT TRANSFM 699 720
FT DOMAIN 721 739
FT DOMAIN 25 105
FT DOMAIN 109 212
FT DOMAIN 223 309
FT DOMAIN 312 399
FT DOMAIN 408 506
FT DOMAIN 511 595
FT DOMAIN 600 684
FT DISULFID 47 95
FT DISULFID 52 99
FT DISULFID 137 195
FT CARBOHYD 273 273
FT CARBOHYD 365 365
FT CARBOHYD 417 417
FT CARBOHYD 463 463
FT CARBOHYD 531 531
FT CARBOHYD 561 561
FT VAAPPLIC 310 402
FT VARIANT 318 318
FT VARIANT 384 384
FT VARIANT 413 413
FT VARIANT 716 716
FT STRAND 26 30
FT STRAND 34 38
FT TURN 39 40
FT STRAND 43 50
FT STRAND 56 61
FT TURN 62 63
FT STRAND 70 74
FT TURN 75 76
FT STRAND 77 82
FT HELIX 87 89
FT STRAND 92 99
FT TURN 100 101
FT STRAND 102 114
FT STRAND 120 123
Query Match 6.7%; Score 155; DB 1; Length 739;
Best Local Similarity 21.9%; Pred. No. 0.004;
Matches 91; Conservative 71; Mismatches 166; Indels 88; Gaps 18;
8 RHLLVLTGALLPAT--QGRKVVIGRGGDVELCTASQKKSIOFHKNKSNQIKILNQ 65
Db ROAVKELOVVISPKNTVIVNPSTKLQGGSVTMTCSSEG.PAPERFW----- 258

```

```

QY 66 GSEFLTKGPSKLANDRADSRRLSDQGNFPLIKKLIKEDSDPTVCE---VEDQKEEVOL 121
Db 259 -----SKKLNGNLQH-----SGNATLTLIMMRBDSGIVYCEGNLIGKRRKEVELI 307
QY 122 V----FGLTANSDTHLIQ--GQSLTLTLESPSSSPVQCRSGKNIQG-----GKT 168
Db 308 VQEKPFVTEISPPRIIAQIGDSVMTLCSVMGCESSPSFSSRTQIDSPLSKRVSEGTNST 367
QY 169 LSVSQLELDQSGTWTCTVLQNKKEFKIDIVLAPQKASIVYK---KEGEVFEFS--- 222
Db 368 LTLSPVSENEHSYLTCTVGTGKKLEKIGLOVELYPPRDEIEMSGGLVNGSSVYTSCKV 427
QY 223 ---FPL-AFTVEKLTSGELMQAERASSKSWITF---DLNKEVSVKVTQDPQLQ- 273
Db 428 PSVYPLDRLEIELKGETTL-----ENIEFLBPDIMKSLKSLMEFTPIFTED 476
QY 274 MGKTL----PLHLTLPOALFOYAGSGNLTALAEAKGKLHOEVNLVYMRATOLQK---N 325
Db 477 TGRALVQQAHLTHIDMEFEFKQROQ---TOTLVYNAVAP--RDTTVLVSPSSILEGSSVN 531
QY 326 LTCEWVGPPSPKMLSLKLENKEAKVSKREKPVVAVNPEAGMMQCLLSDSGVYLE 381
Db 532 WTCLSGCFAPAKTLWSRQLPNEGLOPLSENATLTLSTK-----MEDSGVYLCE 580

RESULT 18
ID NCAM2 HUMAN STANDARD; PRT; 761 AA.
AC P13592; P13593;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neural cell adhesion molecule 1, 120 kDa isoform precursor (N-CAM 120)
DE (NCAM-120) (CD56 antigen).
GN NCAM1 OR NCAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM N-CAM 120).
RC TISSUE=Skeletal muscle;
RX MEDLINE=89305258; PubMed=3253057;
RA Barton C.H., Dickson G., Gower H.J., Rowett L.H., Putt W.,
RA Elsom V., Moore S.E., Goridis C., Walsh F.S.;
RT "Complete sequence and in vitro expression of a tissue-specific
RT phosphatidylinositol-linked N-CAM isoform from skeletal muscle.";
RL Development 104:165-173(1988).
RN [2]
RP SEQUENCE OF 491-761 FROM N.A. (ISOFORM N-CAM 120).
RC TISSUE=Skeletal muscle;
RX MEDLINE=87301755; PubMed=2887295;
RA Dickson G., Gower H.J., Barton C.H., Prentice H.M., Elsom V.L.,
RA Moore S.E., Cox R.D., Quinn C., Putt W., Walsh F.S.;
RT "Human muscle neural cell adhesion molecule (N-CAM): identification
RT of a muscle-specific sequence in the extracellular domain.";
RL Cell 50:1119-1130(1987).
RN [3]
RP SEQUENCE OF 491-655 FROM N.A. (ISOFORM C).
RX MEDLINE=89077552; PubMed=3203385;
RA Gower H.J., Barton C.H., Elsom V.L., Thompson J., Moore S.E.,
RA Dickson G., Walsh F.S.;
RT "Alternative splicing generates a secreted form of N-CAM in muscle
RT and brain.";
RL Cell 55:955-964(1988).
CC -1- FUNCTION: This protein is a cell adhesion molecule involved in
CC neuron-neuron adhesion, neurite fasciculation, outgrowth of
CC neurites, etc.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC Event=Alternative splicing; Named isoforms=3;
CC Name=N-CAM 120;

```

```
CC      IsoId=PI3592-2; Sequence=Displayed;
CC      Name=N-CAM 140;
CC      IsoId=PI3591-1; Sequence=External;
CC      Name=C; Synonyms=Secreted;
CC      IsoId=PI3592-1; Sequence=VSP 002587;
CC      -1 SIMILARITY: Contains 5 immunoglobulin-like C2-type domains .
CC      -1 SIMILARITY: Contains 2 fibronectin type III domains .
CC      -1 DATABASE: NAME=PROW; NOTE=CD guide CD56 entry;
CC      WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd56.htm".
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL collaboration with the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X16841; CA34739.1; -
DR      EMBL; M17409; AAA59912.1; -
DR      EMBL; M22094; AAA59910.1; -
DR      EMBL; M22092; AAA59911.1; -
DR      EMBL; M22091; AAA59911.1; JOINED.
DR      PIR; A31635; A31635.
DR      PIR; S07784; LITHUNG.
DR      Genew; HGNC; 7656; NCAML.
DR      MIM; 116930; -.
DR      GO; GO:0016021; C:integral to membrane; TAS.
DR      GO; GO:0005866; C:plasma membrane; TAS.
DR      InterPro; IPR008957; FN_III-like.
DR      InterPro; IPR003961; FN_III.
DR      InterPro; IPR007110; Ig-I-like.
DR      InterPro; IPR003598; Ig_C2.
DR      Pfam; PF00047; Ig; 5.
DR      SMART; SM00060; FN3; 2.
DR      SMART; SM00408; IGC2; 5.
DR      PROSITE; PS50835; IG_LIKE; 5.
KW      Immunoglobulin domain; Cell adhesion; Glycoprotein; Repeat; Signal;
KV      GPI-anchor; Alternative splicing.
FT      SIGNAL          1..19
FT      CHAIN           20..761
FT                                     NEURAL CELL ADHESION MOLECULE 1, 120 kDa
FT                                     ISOFORM.
FT      DOMAIN          20..111     IG-LIKE C2-TYPE 1.
FT      DOMAIN          116..205     IG-LIKE C2-TYPE 2.
FT      DOMAIN          212..301     IG-LIKE C2-TYPE 3.
FT      DOMAIN          308..403     IG-LIKE C2-TYPE 4.
FT      DOMAIN          406..491     IG-LIKE C2-TYPE 5.
FT      DOMAIN          518..595     FIBRONECTIN TYPE-III 1.
FT      DOMAIN          660..727     FIBRONECTIN TYPE-III 2.
FT      DISULFID        41..96       PROBABLE.
FT      DISULFID        139..189     PROBABLE.
FT      DISULFID        235..287     PROBABLE.
FT      DISULFID        329..385     PROBABLE.
FT      DISULFID        426..479     PROBABLE.
FT      CARBOHYD        222..222     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD        315..315     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD        347..347     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD        423..423     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD        449..449     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD        478..478     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      VARSPPLIC       635..655     GEPSPAPKLEGQWDGNGNSIKV -> NIAGNHCNMFOAGLH
FT                                     NALMK (In isoform C).
FT                                     /FTIG=VSP 002587.
SQ      SEQUENCE       761 AA;   83770 MW;   FPCADJ252D7A867E CR664;
Query Match      6.2%; Score 145; DB 1; Length 761;
Beet Local Similarity 20.5%; Pred. No. 0.02;
Matches 100; Conservative 67; Mismatches 211; Indels 110; Gaps 20;
3 RG-VFRRLHLVLQLALPAATQGKKVVLGGKGDTVELTCTASQKKSLTFHHKSNQRI 61
196 RGEINFKIQIVLVNPPTTGAAQNITVNATNVANISGVTVTCADGEGPEPTMSV----- 247
```

| QY | 62 | LGNGSFLTKRGSKLANDRADSRSLMDGNPLIIKXNKLTEDSPTYCEVE---- | DOKEE | 116 |
|-------------|--|--|------------|---------|
| Db | 248 | -----TKDGSQIQEEDDEKTYIFSDSSQLTIKKVDKNDEARYICIAENKAGEODAT | 299 | |
| QY | 118 | VOLLVEG-----LPASDTHLLQGSGLTLTLESPPGSSPSVQCR----- | SPRKNIOGQ- | 166 |
| Db | 300 | IIHKVFAKKRTIYVENQGTAMELEBQVTLTCEASGDIPLSIIMRSTRNISSEKTLIDGHM | 359 | |
| QY | 167 | -----KTLVSQLELQDSSGTMTCTYVQ--NQKVEFKIDIVLAFQKASSIVYKKEG | 216 | |
| Db | 360 | VVRSARVSLTLTKSIQYTDAGEYICTASNTIGDOSQSMYLEVQYAPKLGQPVAVVWMEG | 419 | |
| QY | 217 | EYGE-----FSPPLAFTVEKLTGSELMMQABRSASSKSWITPDLK--KEVSVRVYQDP | 270 | |
| Db | 420 | NQVNTICEVFAIPSA-TIS-----WFRDQQLPSNSNINIKYITPESASILEVTPDS | 470 | |
| QY | 271 | KLOMGK-----KLPLHLTLPOA-----LPOYAGSGLTLLEAKTKLHQ | 310 | |
| Db | 471 | ENDFGVYNGCTAVMIRIGQESLEPIIVQADTPSPSIDQEPVSSPAQVQDPDEPATG---- | 526 | |
| QY | 311 | EVNLVVMRATQLOKXNLTCGWCPTSPKMLSLKLENK--EAKYSKREKPVV--LNPQAG | 366 | |
| Db | 527 | --GVPIITKKAEMBRVAGEVW-----HSKRYDKKESMSGIVTIYGLKXETT | 571 | |
| QY | 367 | MMQCLSDSGOVLLE-----SNIKVLPWTSTVHPASALPAPPGSALPDPQRTASALPDP | 422 | |
| Db | 572 | YAVRLAALNGKLGIGELISAASEFKTQPVHSP--PPASASSSTVPVLPSPD----- | TTWPLP | 625 |
| QY | 423 | PAASALPA 430 | | |
| Db | 626 | ALATTEPA 633 | | |
| RESULT 19 | | | | |
| VCAM1 MOUSE | | | | |
| ID | VCAM1 MOUSE | STANDARD; | PRT; | 739 AA. |
| AC | P29533; | | | |
| DT | 01-APR-1993 (Rel. 25, Created) | | | |
| DT | 01-APR-1993 (Rel. 25, Last sequence update) | | | |
| DT | 10-OCT-2003 (Rel. 42, Last annotation update) | | | |
| DE | Vascular cell adhesion protein 1 precursor (V-CAM 1). | | | |
| GN | VCAM1 OR VCAM-1. | | | |
| OS | Mus musculus (Mouse). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| OX | NCBI_TaxID=10090; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 1). | | | |
| RC | STRAIN=FVB; TISSUE=Lung; | | | |
| RC | MEDLINE=92161437; PubMed=1371918; | | | |
| RX | Heesien C., Moy P., Tizard R., Chisholm P., Williams C., Wyse M., | | | |
| RA | Burkly L., Miyake K., Kincaid P., Lobb R.; | | | |
| RT | "Cloning of murine and rat vascular cell adhesion molecule-1."; | | | |
| RL | Biochem. Biophys. Res. Commun. 183:163-169(1992). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 1). | | | |
| RC | TISSUE=Lymph node; | | | |
| RC | MEDLINE=93246254; PubMed=7683304; | | | |
| RX | Araki M., Araki K., Vassalli P.; | | | |
| RA | "Cloning and sequencing of mouse VCAM-1 cDNA."; | | | |
| RL | Gene 126:261-264(1993). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 1). | | | |
| RC | STRAIN=129; TISSUE=Embryo; | | | |
| RC | MEDLINE=94117008; PubMed=7507076; | | | |
| RA | Cybulsky M.I., Allan-Motamed M., Colline T.; | | | |
| RT | "Structure of the murine VCAM1 gene."; | | | |
| RL | Genomics 18:387-391(1993). | | | |
| RN | [4] | | | |
| RP | SEQUENCE OF 1-693 FROM N.A. (ISOFORM 1). | | | |
| RC | STRAIN=NH Swiss, and 129/Sv; | | | |
| RC | Kumar A.G., Dai Y.X., Kozak C.A., Mims M.P., Gotto A.M. Jr., | | | |


```

Cc Db Qy Db Qy Db Qy Db Qy Db Qy DB RESULT_20
Cc Db Qy Db Qy Db Qy DB NEOI_MOUSE
Cc AC ID -NEOI_MOUSE STANDARD; PRT; 1493 AA. p97798;
Cc DT 16-OCT-2001 (Rel. 40, Created)
Cc DT 16-OCT-2001 (Rel. 40, Last sequence update)
Cc DT 10-OCT-2003 (Rel. 42, Last annotation update)
Cc DE Neogenin precursor.
Cc GN NEOL OR NGN.
Cc OS Mus musculus (Mouse).
Cc OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Cc OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Cc CX NCBI_TaxID=10090;

Cc RN [1]
Cc RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
Cc RC TISSUE=BRAIN;
Cc RX MEDLINE=97407661; Pubmed=9264410;
Cc RA Keeling S.L., Gad J.M., Cooper H.M.;
Cc RT "House neogenin, a DCC-like molecule, has four splice variants and is
Cc RL expressed widely in the adult mouse and during embryogenesis.";
Cc RT Oncogene 15:691-700(1997).
Cc CC -I FUNCTION: May be involved as a regulatory protein in the
Cc CC transition of undifferentiated proliferating cells to their
Cc CC differentiated state. May also function as a cell adhesion
Cc CC molecule in a broad spectrum of embryonic and adult tissues.
Cc CC -I SUBCELLULAR LOCATION: Type I membrane protein.
Cc CC -I ALTERNATIVE PRODUCTS:
Cc CC Event=Alternative splicing; Named isoforms=5;
Cc CC Comment=Additional isoforms seem to exist;
Cc CC Name=1; IsoId=P97798-1; Sequence=Displayed;
Cc CC Name=2; IsoId=P97798-2; Sequence=VSP_002594;
Cc CC Name=3; IsoId=P97798-3; Sequence=VSP_002595;
Cc CC Note=Expression developmentally regulated;
Cc CC Name=4; IsoId=P97798-4; Sequence=VSP_002596;
Cc CC Note=Expression developmentally regulated;
Cc CC Name=5; IsoId=P97798-5; Sequence=VSP_002597;
Cc CC Note=Expression developmentally regulated;
Cc CC -I TISSUE SPECIFICITY: Widely expressed.
Cc CC -I DEVELOPMENTAL STAGE: EXPRESSED UBICITOUSLY THROUGHOUT THE MID TO
Cc CC LATE STAGES OF GESTATION AND IN ADULT TISSUES. STRONG EXPRESSION
Cc CC IS OBSERVED IN THE VENTRAL REGION OF THE VENTRICULAR ZONE OF THE
Cc CC E15.5 MOUSE NEURAL TUBE, AS WELL AS IN THE VENTRICULAR ZONES OF
Cc CC THE MESENCEPHALON AND RHOMBENCEPHALON. ISOFORMS 3 AND 4 ARE
Cc CC EXPRESSED AT HIGHER LEVEL COMPARED TO OTHER ISOFORMS BETWEEN E11.5
Cc CC AND E16.5.
```

| | | |
|--|--|--|
| CC | -I- SIMILARITY: | Belongs to the immunoglobulin superfamily, DCC family. |
| CC | -I- SIMILARITY: | Contains 4 immunoglobulin-like C2-type domains. |
| CC | -I- SIMILARITY: | Contains 6 fibronectin type III domains. |
| CC | ----- | |
| CC | This SWISS-PROT entry is copyright. | It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). |
| CC | ----- | |
| DR | EMBL; Y09535; CAAT70727.1; .- | |
| DR | HSSP; P02751; 1YTF. | |
| DR | MED; MG1.1097159; Neol. | |
| DR | InterPro; IPRO08957; FN_III-like. | |
| DR | InterPro; IPRO03961; FN_III. | |
| DR | InterPro; IPRO03962; FNIII subd. | |
| DR | InterPro; IPRO07110; Ig-like. | |
| DR | InterPro; IPRO03598; Ig_C2. | |
| DR | Pfam; PF00047; fn3; 6. | |
| DR | Pfam; PF00047; Ig; 4. | |
| DR | PRINTS; PR00014; ENTPEIIII. | |
| DR | SMART; SMO0060; FN3; 6. | |
| DR | SMART; SMO0408; IGc2; 4. | |
| DR | PROSITE; PS50835; Regat; Signal; Transmembrane; Immunoglobulin domain; | |
| KV | Cell adhesion; Repeat; Signal; Signal; Alternative splicing. | |
| KW | Glycoprotein; Alternative splicing. | |
| FT | SIGNAL | 1 36 POTENTIAL. |
| FT | CHAIN | 37 1493 NEOGENIN. |
| FT | DOMAIN | 37 1136 EXTRACELLULAR (POTENTIAL). |
| FT | TRANSNEM | 1137 1157 POTENTIAL. |
| FT | DOMAIN | 1158 1493 CYTOPLASMIC (POTENTIAL). |
| FT | DOMAIN | 63 158 IG-LIKE C2-TYPE 1. |
| FT | DOMAIN | 163 249 IG-LIKE C2-TYPE 2. |
| FT | DOMAIN | 254 347 IG-LIKE C2-TYPE 3. |
| FT | DOMAIN | 352 437 IG-LIKE C2-TYPE 4. |
| FT | DOMAIN | 467 564 FIBRONECTIN TYPE-II 1. |
| FT | DOMAIN | 567 660 FIBRONECTIN TYPE-II 2. |
| FT | DOMAIN | 661 760 FIBRONECTIN TYPE-II 3. |
| FT | DOMAIN | 766 860 FIBRONECTIN TYPE-II 4. |
| FT | DOMAIN | 881 981 FIBRONECTIN TYPE-III 5. |
| FT | DOMAIN | 982 1083 FIBRONECTIN TYPE-III 6. |
| FT | DOMAIN | 1149 1153 POLY-VAL. |
| FT | DISULFID | 85 140 BY SIMILARITY. |
| FT | DISULFID | 184 232 BY SIMILARITY. |
| FT | DISULFID | 281 331 BY SIMILARITY. |
| FT | DISULFID | 373 421 BY SIMILARITY. |
| FT | CARBOHYD | 84 84 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 221 221 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 337 337 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 501 501 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 520 520 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 670 670 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 746 746 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD | 940 940 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | VARSPLIC | 442 461 Missing (in isoform 2). |
| FT | VARSPLIC | /FTId-VSP 002594. |
| FT | VARSPLIC | Missing (in isoform 3). |
| FT | VARSPLIC | /FTId-VSP 002595. |
| FT | VARSPLIC | Missing (in isoform 4). |
| FT | VARSPLIC | /FTId-VSP 002596. |
| FT | VARSPLIC | Missing (in isoform 5). |
| FT | VARSPLIC | /FTId-VSP 002597. |
| SO | SEQUENCE | 1493 AA; 163159 MW; 441DE919D5E17COB CRC64; |
| Query Match | | |
| Best Local Similarity | 6.1%; Score 141.5; DB 1; Length 1493; | |
| Matches 117; Conservative 54; Mismatches 188; Indels 145; Gaps 28; | | |
| DQ | 10 LLLVLQALL--PA-----TGKNV-----VLGKKDTVEIT 40 | |
| | : | |
| DB | 25 LLLLLLLLPLLLIGDPASGAATKSGPRGSGASVRFTFFFLVIVBPVDLTSVRGSSVILN 84 | |

QY 41 CTASQKSIQFHMKNSQIKILGNGSFLLTKGPSRLRADRSRLMDQGNFPLIKL- 99
 DB 85 CSAVSEPEPNIEWK-----KQGT-----LNIESDRQLLPDGS--LFISNV 126
 QY 100 ----KIEDSDYIC--EVED-----QKEEVQLLVFGL----TASDTHLLQGQSLTLTLES 145
 DB 127 HSKNNKPEGEFYOCATVATNIGTTYSRTAKLTAVGLPFTSQPEPSVYVGNAILNCEV 186
 QY 146 PRGSSPV---QCRRP-----RGKNIQGGKTLVSQLELOSGTWTCTVLQN-----QKK 192
 DB 187 NADLVFPRMEONRPLLDDIVKLPSG-TLVISNATEGGGLRCLVSGSGPPKSE 245
 QY 193 VEFKI-----DIVVAFQKASSIVYKKEGVEFSFPLAFYKLTGSGELMWQERASS 247
 DB 246 AELKVLQPEEIVDLVFLMRPSMMKVYQ-----SAVLPVSGLPVAVW----- 293
 QY 248 SKSWITPDLKNKEVSVKAVTQDPKLOMGKKLPLHLTLPLQALPQY--AGSGNLT----- 298
 DB 294 -----MKNEEVLDTSSGRLVLAGGCLISDVEDDAGTYFCIADNGNKTVEAQAE 345
 QY 299 LALEAKTKLQHOEVLVVMRATOLQKILTCFVWGPTSKML-----SLKLENKAKVSKR 354
 DB 346 LTVQVPPEFLQOPANIVAHESMDIV--FECEVTGKPTVVMKXGADVIPSDFKIVKE 403
 QY 355 E--KPVVILNPEAGMGOCLL--SDSGQ-----VLLESNIKVLPTWSTPVHPRASALPAP 404
 DB 404 HNLQVLGLVKGDEGFGYQCIENDVGNAGQMLILEHDV-AIPL-----LP 449
 QY 405 PTG--SALPD---PQTASALPDP 423
 DB 450 PTLSTSATTDHLAPATGTPPLPAP 473

RESULT 21

SNE2_HUMAN STANDARD; PRT; 6885 AA.
 AC Q8MXH0; Q8N1S3; Q8NF49; Q8TER7; Q8MWM3; Q8MWM4; Q8MWM5; Q8MXH1;
 AC Q9NU50; Q9UFQ4; Q9Y2L4; Q9Y4R1;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neepirin 2 (Nuclear envelope spectrin repeat protein 2) (Syn-2)
 DE (Synaptic nuclear envelope protein 2) (Nucleus and actin connecting
 element protein) (NUANCE protein).
 GN SNE2 OR NUA OR KIAA1011.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 9), FUNCTION, CHARACTERIZATION, AND
 RP INTERACTION WITH F-ACTIN.
 RX MEDLINE=22113122; PubMed=12118075;
 RA Zhen Y.-Y., Libotte T., Munck M., Noegel A.A., Korenbaum E.;
 RT "NUANCE, a giant protein connecting the nucleus and actin
 RT cyoskeleton.";
 RL J. Cell Sci. 115:3207-3222(2002).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 4; 5 AND 7).
 RX MEDLINE=21652858; PubMed=11792814;
 RA Zhang Q., Skepper J.N., Yang F., Davies J.D., Hegyi L., Roberts R.G.,
 RA Weisberg P.L., Ellis J.A., Shanahan C.M.;
 RT "Neepirins: a novel family of spectrin-repeat-containing proteins that
 RT localize to the nuclear membrane in multiple tissues.";
 RL J. Cell Sci. 114:4485-4498(2001).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=22296983; PubMed=12408964;
 RA Zhang Q., Ragnauth C., Greener M.J., Shanahan C.M., Roberts R.G.;
 RT "The neepirins are giant actin-binding proteins, orthologous to
 RT Drosophila melanogaster muscle protein MSP-300.";
 RL Genomics 80:473-481(2002).

RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 6).
 RC TISSUE=Testis;
 RA Pousset A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX PubMed=12508121;
 RA Hellig R., Beckenberg R., Petit J.-L., Fonknechten N., Da Silva C.,
 RA Cattelico L., Levy M., Barbe V., De Berardinis V., Ureta-Vidal A.,
 RA Pelletier E., Vico V., Anthonard V., Rowen L., Madan A., Qin S.,
 RA Sun H., Du H., Pepin K., Artiguenave F., Robert C., Cruaud C.,
 RA Briels T., Jallou O., Friedlander L., Sanson G., Brothier P.,
 RA Cure S., Segures B., Aniere F., Samain S., Cregeau H., Abbs N.,
 RA Alich N., Boscus D., Dickhoff R., Dors M., Dubois I., Friedman C.,
 RA Gouyenoux M., James R., Madan A., Mailey-Bstrada B., Mangelot S.,
 RA Martins N., Menard M., Ozas S., Ratcliffe A., Shaffer T., Traak B.,
 RA Vacherie B., Bellemere C., Belser C., Beshard-Gonnet M.,
 RA Battoi-Mavel D., Boutard M., Briez-Silla S., Combette S.,
 RA Dufosse-Laurent V., Ferron C., Lechaplais C., Louesse S.,
 RA Magdeleat G., Patrau E., Petit E., Sirvain-Trukiewicz P., Trybou A.,
 RA Vega-Czarny N., Bataille E., Bluet E., Bodelais I., Dubois M.,
 RA Dumont C., Guerin T., Hatray S., Hammadi R., Muanga J., Pellouin V.,
 RA Robert D., Wunderle E., Gauguier G., Roy A., Sainte-Marthe L.,
 RA Verdier J., Verdier-Discalca C., Hillier L.W., Fulton L., McPherson J.,
 RA Matsuda F., Wilson R., Scarpelli C., Gyapay G., Winkler P., Saurin W.,
 RA Quetier F., Waterson R., Hood L., Weissbach J.;
 RT "The DNA sequence and analysis of human chromosome 14.";
 RL Nature 421:601-607(2003).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 9).
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hale F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toehiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullah S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 1-956 AND 5133-6885
 RP FROM N.A.
 RC TISSUE=Spleen, and Tongue;
 RA Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.,
 RA Nimomiyu K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
 RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
 RA Katsuya N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
 RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahara K., Masuno Y., Nagai K., Isogai T.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE OF 5709-6885 FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RX MEDLINE=99246063; PubMed=10231032;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirogawa M.,

```

RA  Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT  "Prediction of the coding sequences of unidentified human genes. XIII
RT  The complete sequences of 100 new cDNA clones from human which code
RT  for large proteins in vitro.";
RL  DNA Res. 6:63-70(1999).
RN  [9]
RP  REVISIONS.
RC  TISSUE=Brain.
RX  MEDLINE=22158693; PubMed=12168954;
RA  Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT  "Construction of expression-ready cDNA clones for KIAA genes: manual
RT  curation of 330 KIAA cDNA clones.";
RL  DNA Res. 9:99-106(2002).
RN  [10]
RP  SEQUENCE OF 5754-6885 FROM N.A.
RX  MEDLINE=21154917; PubMed=11230166;
RA  Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glasil S.,
RA  Ansorge W., Boecker M., Bloeker H., Baerzachs S., Blum H.,
RA  Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strick N.,
RA  Mewes H.-W., Otenwaelder B., Obernair B., Tampe J., Heubner D.,
RA  Mambert R., Korn B., Klein M., Poustka A.;
RT  "Towards a catalog of human genes and proteins: sequencing and
RT  analysis of 500 novel complete protein coding human cDNAs.";
RL  Genome Res. 11:422-435(2001).
CC  -1 FUNCTION: Involved in the maintenance of nuclear organization and
CC  structural integrity. Probable anchoring protein which tethers the
CC  nucleus to the cytoskeleton. Connects nuclei to the cytoskeleton
CC  by interacting with the nuclear envelope and with F-actin in the
CC  cytoplasm.
CC  -1 SUBUNIT: Interacts with F-actin via its N-terminal domain.
CC  -1 SUBCELLULAR LOCATION: Type IV membrane protein (Potential). The
CC  largest part of the protein is cytoplasmic, while its C-terminal
CC  part is associated with the nuclear envelope, most probably the
CC  outer nuclear membrane. Remains associated with the nuclear
CC  envelope during its breakdown in mitotic cells.
CC  -1 ALTERNATIVE PRODUCTS:
CC  Event=Alternative splicing; Named isoforms=9;
CC  Name=1;
CC  IsoId=Q8WXH0-1; Sequence=Displayed;
CC  Name=2;
CC  IsoId=Q8WXH0-2; Sequence=VSP_007164, VSP_007166;
CC  Note=No experimental confirmation available;
CC  Name=3;
CC  IsoId=Q8WXH0-3; Sequence=VSP_007155;
CC  Note=Produced by exon skipping that results in a frameshift. No
CC  experimental confirmation available;
CC  Name=4; Synonyms=Beta;
CC  IsoId=Q8WXH0-4; Sequence=VSP_007156;
CC  Name=5; Synonyms=Alpha;
CC  IsoId=Q8WXH0-5; Sequence=VSP_007157, VSP_007164, VSP_007165;
CC  Name=6;
CC  IsoId=Q8WXH0-6; Sequence=VSP_007158, VSP_007165, VSP_007166;
CC  Note=No experimental confirmation available;
CC  Name=7; Synonyms=Gamma;
CC  IsoId=Q8WXH0-7; Sequence=VSP_007154, VSP_007163;
CC  Name=8;
CC  IsoId=Q8WXH0-8; Sequence=VSP_007161, VSP_007162;
CC  Note=No experimental confirmation available;
CC  Name=9; Synonyms=NUANCE-N-33;
CC  IsoId=Q8WXH0-9; Sequence=VSP_007159, VSP_007160;
CC  -1 TISSUE SPECIFICITY: Widely expressed, with higher level in kidney
CC  adult and fetal liver, stomach and placenta. Weakly expressed in
CC  skeletal muscle and brain. Isoform 5 is highly expressed in
CC  pancreas, skeletal muscle and heart.
CC  -1 DOMAIN: The Klarsicht domain mediates the nuclear envelope
CC  targeting.
CC  -1 SIMILARITY: Belongs to the negrin family.
CC  -1 SIMILARITY: Contains 1 actin-binding domain.
CC  -1 SIMILARITY: Contains 2 calponin-homology (CH) domains.
CC  -1 SIMILARITY: Contains 1 Klarsicht domain.
CC  -1 SIMILARITY: Contains 1 leucine-rich (LRR) repeats.
CC  -1 SIMILARITY: Contains 9 spectrin repeats.

```

```

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation on its
CC the European Bioinformatics Institute. There are no restrictions on the use
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.embnet.ch/announce/
CC or send an email to license@ib5b.ch).
CC -----
DR EMBL, AF435010, AALJ3547.1, -
DR EMBL, AF435011, AALJ3548.1, -
DR EMBL, AY061757, AALJ3800.1, -
DR EMBL, AY061758, AALJ3801.1, -
DR EMBL, AY061759, AALJ3802.1, -
DR EMBL, AF495911, AAN60443.1, -
DR EMBL, AL117404, CAB55905.1, -
DR EMBL, AL162832, -; NOT_ANNOTATED CDS.
DR EMBL, ALJ35094, -; NOT_ANNOTATED CDS.
CC -----
Query Match 6.1%; Score 141; DB 1; Length 6885;
Best Local Similarity 25.6%; Pred. No. 0.65; Mismatches 128; Indels 54; Gaps 15;
Matches 83; Conservative 59;
QY 96 IKNIKIEDSPYICEVEDQKEEVQVLVFGTLA--NSDTLLIQ--GQSLTTLTLESPPGSSP 151
Db 592 VENVRL---LRACFEETKKKEIKEVEFETLACQNLNLSHATLNENAGNFLVEVSNDDVGSISI 647
QY 152 SVQGRSPRGNIQGGKTLTSSQLEL-----QDSGTW--TCVVLQNKQ--VEPKIDI 199
Db 648 SKELRR--LKKRRKRVKSVSKTQLEMLPLMIKKODQPTFDNSGNILSKKEKATVEFTSDM 704
QY 200 VVLAFOKASSIV-----YKKEGEQVEFSFPLAFVETKLTSGELMWOAERASSKSWITF 254
Db 705 SVELPENYQNQNIKAGEHKEKNEEFQTQLVAKAOKVEKLTIGQVEI--WEAE---AKSVLDQ 759
QY 255 D--LAKKEVGVKRV---TDDPKLQMGKRLPLHLTFQALPQVAGSGNLTLLALEAKTGKL 308
Db 760 DDVDTSMEEELKHLIAKGSWFDELMARSEDMLOMDIONISSQESFQVHLTTGLOAKIQEA 819
QY 309 HOEVLVVMVATQLOKRLTCEVWGPSTSPKMLSLKLENKEAVSKREKPVVVLNPEAGMW 368
Db 820 KEKQIVNVVLIALAKVLT-----DVSPDDIRLKMBE-----SQKLESTYMAAQQLLG 869
QY 369 QC-----LLSDGQVLLSNNIKVL 387
Db 870 QRESFGELISKHKEALIIISNTKSL 893
CC -----
RESULT 22
DCC_MOUSE STANDARD; PRT; 1447 AA.
AC P70211;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tumor suppressor protein DCC precursors.
GN DCC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10099;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A, B AND C).
RC STRAIN=BALB/c; TISSUE=Brain;
RC MEDLINE=96112625; PubMed=8570174;
RA Cooper H.M., Armes P., Britto J., Gad J., Wilks A.F.,
RT "Cloning of the mouse homologue of the deleted in colorectal cancer
RT gene (mDC) and its expression in the developing mouse embryo.";
RL Oncogene 11:2243-2254(1995).
RN [2]
RN REVISIONS.
RC STRAIN=BALB/c; TISSUE=Brain;
RC Cooper H.M.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

```

```

CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=A;
CC IsoId=P70211-1; Sequence=Displayed;
CC Note=Isoform B is produced by alternative initiation at Met-85
CC of isoform A;
CC Name=C;
CC IsoId=P70211-2; Sequence=VSP_002501;
CC Event-Alternative initiation;
CC Comment=2 isoforms, A (shown here) and B, are produced by
CC alternative initiation at Met-1 and Met-85;
CC -1- TISSUE SPECIFICITY: In the embryo, expressed at high levels in the
CC developing brain and neural tube. In adult, highly expressed in
CC brain with very low levels found in testis, heart and thymus.
CC Isoform C is expressed only in the embryo.
CC -1- DEVELOPMENTAL STAGE: Low levels in early gestation. Highest levels
CC expressed during mid gestation. Levels decrease in late gestation
CC and remain at this level in the adult.
CC -1- SIMILARITY: Belongs to the immunoglobulin-like C2-type domain.
CC -1- SIMILARITY: Contains 6 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; X85788; CAA59786.1; -.
DR HSPB; P56276; ITLK.
DR MGD; MGI:94869; DCC.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003962; FnIII subd.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00047; Ig; 4.
DR PRINTS; PR00014; FNTYPEIII.
DR SMART; SMO0060; FN3; 6.
DR SMART; SMO0408; IGc2; 3.
DR PROSITE; PS50835; IG-LIKE; 4.
KM Glycoprotein; Immunoglobulin domain; Transmembrane; Signal;
KM Repeat; Anti-oncogene; Alternative initiation; Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 1447 TUMOR SUPPRESSOR PROTEIN DCC, ISOFORM A.
FT INIT MET 85 85 FOR ISOFORM B.
FT DOMAIN 26 1097 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1098 1122 POTENTIAL.
FT DOMAIN 1123 1447 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 136 135 IG-LIKE C2-TYPE 1.
FT DOMAIN 139 229 IG-LIKE C2-TYPE 2.
FT DOMAIN 234 326 IG-LIKE C2-TYPE 3.
FT DOMAIN 331 416 IG-LIKE C2-TYPE 4.
FT DOMAIN 426 522 FIBRONECTIN TYPE-III 1.
FT DOMAIN 525 618 FIBRONECTIN TYPE-III 2.
FT DOMAIN 619 716 FIBRONECTIN TYPE-III 3.
FT DOMAIN 722 816 FIBRONECTIN TYPE-III 4.
FT DOMAIN 840 940 FIBRONECTIN TYPE-III 5.
FT DOMAIN 941 1042 FIBRONECTIN TYPE-III 6.
FT DISULFID 61 117 BY SIMILARITY.
FT DISULFID 161 212 BY SIMILARITY.
FT DISULFID 261 310 BY SIMILARITY.
FT DISULFID 352 400 BY SIMILARITY.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 702 702 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPIC 819 838 Missing (in isoform C).
SQ SEQUENCE 1447 AA; 158298 MW; 0D1F1097C2D25B9F CRC64;
Query Match
Best Local Similarity 20.9%; Pred. No. 0.12;
Matches 76; Conservative 48; Mismatches 131; Indels 108; Gaps 14;
6.0%; Score 139; DB 1; Length 1447;
QY 34 GDTVELTCTASQKSIQFMKNSNQIKLGNQSPFTKGPSKLNDRADSRRLMDQGNF 92
DB 154 GDTVLTKCEVIGBPMFTIHWQKNOQ-----DINPLPGDSRVVLPBG-- 195
QY 93 PLIKLKLDSDPTTYICEVD-----QKEVQL-----VRGLTANSDTHLQ 137
DB 196 ALQISRLQPDSDSVYSCSARNPASIRTGNAEVRITSDPGLHRLQVFLQRPNSVIALEGK 255
QY 138 SLTL---TLESPPGS-----SPSVOCSPRGKNIQSGKTLSVQLELQSGTWTCTVLQ 188
DB 256 DAYLECCVSGYPPPSFTWLRGEEVQLRSKQ-YSLGSGWLLISNVTDDSDGTYTCVY 314
QY 189 NQKKEFKIDIVLA---FQKASSIYKKEGEVPSFPLAFTVEKLTGSGELMWAQERA 245
DB 315 KNNISASAEITVLPWPMFLNHPNSNLVAYESMDIEF-----ECA 353
QY 246 SSSKSWTTP-LNKKEVSYSRVQDPKLGKPLHLTLPOALPOYAGSGNLT----- 299
DB 354 VSGKPVFTVMKNGDV-----VPSYFQIVGSSNRLILGVK 392
QY 300 -----ALEATGKLGHEVNLVWRAVATQLQKLTCEVGPSPKMLSLKLENKAV 351
DB 393 SDGFPYCVAVENAGNAGNSAQLIVPKPAIPSSIL-----PSAPHDVLPVLSRVRL 447
QY 352 SKR 354
DB 448 SWR 450
RESULT 23
NCM2_MOUSE STANDARD; PRT; 837 AA.
ID NCM2_MOUSE
AC O35136; O35962;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neutral cell adhesion molecule 2 precursor (N-CAM 2) (RB-8 neural cell
DE adhesion molecule) (R4B12).
GN NCM2 OR OCAM OR RNCAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RC STRAIN=BALB/c; TISSUE=Olfactory neuroepithelium;
RX MEDLINE=97368238; Pubmed=9221781;
RA Yoshitara Y., Kawasaki M., Tamada A., Fujita H., Hayashi H.,
RA Kagamiyama H., Mori K.;
RT "OCAM: A new member of the neural cell adhesion molecule family
RT related to zone-to-zone projection of olfactory and vomeronasal
RT axons.";
RL J. Neurosci. 17:5830-5842(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RC STRAIN=C57BL/6J; TISSUE=Olfactory epithelium;
RX MEDLINE=97476194; Pubmed=9334170;
RA Alenius M., Bohm S.;
RT "Identification of a novel neural cell adhesion molecule-related gene
RT with a potential role in selective axonal projection.";
RL J. Biol. Chem. 272:26083-26086(1997).
CC -1- FUNCTION: May play important roles in selective fasciculation and

```

| CC | zone-to-zone projection of the primary olfactory axons. |
|----|--|
| CC | -1- SUBCELLULAR LOCATION: Type I membrane protein (long isoform) and attached to the membrane by a GPI-anchor (short isoform). |
| CC | -1- ALTERNATIVE PRODUCTS: |
| CC | Event-Alternative splicing; Named isoforms=2; |
| CC | Name:Long; |
| CC | Idoid=035136-1; Sequence=Displayed; |
| CC | Name:Short; |
| CC | Idoid=035136-2; Sequence=VSP_002590; |
| CC | -1- TISSUE SPECIFICITY: Expressed in subsets of both olfactory and vomeronasal neurons in a zone-specific manner. |
| CC | -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains. |
| CC | -1- SIMILARITY: Contains 2 fibronectin type III domains. |
| CC | ----- |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-stb.ch/announce or send an email to license@isb-stb.ch). |
| CC | ----- |
| DR | EMBL; AF001287; AAB69125.1; - |
| DR | EMBL; AF001286; AAB69124.1; - |
| DR | EMBL; AF016619; AAC53375.1; - |
| DR | MGD; WGI:97282; Ncam2. |
| DR | InterPro; IPR008957; FN_III-like. |
| DR | InterPro; IPR003961; FN_III. |
| DR | InterPro; IPR007110; IG_1-like. |
| DR | InterPro; IPR003598; IG_C2. |
| DR | Pfam; PF00041; fn3; 2. |
| DR | Pfam; PF00047; Ig; 5. |
| DR | SMART; SM00060; FN3; 2. |
| DR | SMART; SM00408; IGc2; 5. |
| DR | PROSITE; PS00835; IG_LIKE; 5. |
| KW | Cell adhesion; Transmembrane; |
| KW | Immunoglobulin domain; Signal; GPI-anchor; Alternative splicing. |
| KW | SIGNAL |
| FT | 1 |
| FT | CHAIN |
| FT | 20 837 |
| FT | DOMAIN |
| FT | 20 697 |
| FT | TRANSMEM |
| FT | 698 718 |
| FT | DOMAIN |
| FT | 719 837 |
| FT | DOMAIN |
| FT | 21 108 |
| FT | DOMAIN |
| FT | 113 202 |
| FT | DOMAIN |
| FT | 208 297 |
| FT | DOMAIN |
| FT | 302 396 |
| FT | DOMAIN |
| FT | 401 491 |
| FT | DOMAIN |
| FT | 482 581 |
| FT | DOMAIN |
| FT | 594 678 |
| FT | DISULFID |
| FT | 42 93 |
| FT | DISULFID |
| FT | 136 186 |
| FT | DISULFID |
| FT | 232 281 |
| FT | DISULFID |
| FT | 322 380 |
| FT | DISULFID |
| FT | 422 475 |
| FT | CARBOHYD |
| FT | 177 177 |
| FT | CARBOHYD |
| FT | 219 219 |
| FT | CARBOHYD |
| FT | 309 309 |
| FT | CARBOHYD |
| FT | 406 406 |
| FT | CARBOHYD |
| FT | 419 419 |
| FT | CARBOHYD |
| FT | 445 445 |
| FT | CARBOHYD |
| FT | 474 474 |
| FT | CARBOHYD |
| FT | 562 562 |
| FT | CARBOHYD |
| FT | 694 837 |
| FT | VARSPLIC |
| FT | 694 837 |
| FT | SEQUENCE |
| FT | 837 AA; 93203 MW; 704738053A2D65A5 CRC64; |
| FT | ----- |
| FT | Query Match |
| FT | Score 138; DB 1; Length 837; |
| FT | Match Local Similarity 19.2%; Pred. No. 0.06; |
| FT | Matches 93; Conservative 74; Mismatches 190; Indels 128; Gaps 22; |

| | | | | |
|------------|---|-----------------------------------|---|---------|
| OY | | 3 | GG-VERRLLLVQLALLPAATQGVKVLGKKGDPTVELCTPSQOKSIQIFHMKNSQIKI | 61 |
| Dd | | 193 | KGEIDFRDIIVAVNVPRAIMPQKSFMAATEAGBEETLCKKSGSPDPTISWFRNG--KL | 250 |
| OY | | 62 | LGNQGSPLTKGPSKLNDRASRRSLMDQGNFPLLIIKNLEIEDPTYICEV----EDQKE | 116 |
| Dd | | 251 | IENNKYLKG-----SNTSLTYRNIINMGOGSYVCATATNAAGBDQX | 293 |
| OY | | 117 | E-VQLLVGLTANSPTHLQGOS-----LTTLSEPPSGSSPSVQC----- | 155 |
| Dd | | 294 | AFLQVFV-----OPHILQLNETTSENGHTLVLCAGEBEPVEITWRRAIDGVMFSEG | 346 |
| OY | | 156 | -RSPPG-----KNIOGKITLSVSQLELDSDGTTCVTLQ-----NOKKVERKIDIVLAFOK | 206 |
| Dd | | 347 | DSPEGRILEVKQHQRSSLDHINDVCLDSGRDYDCBASRTIGGHQRMHLDIETAP-KFVS | 405 |
| OY | | 207 | ASSIYYKKEGEGEVESFPFLAFTVEKLTGSGELMOWERAASSKSXWTIPDLKNNEVASKVY | 266 |
| Dd | | 406 | NQMYYSMEGNPINISCVT-----ANPPASJHMREKRL-----LPKQT | 446 |
| OY | | 267 | TODPELQNGKKULPHLTLPOALPOYAGSNLTIALAEAKTGKHOEYNLYVMRATOLOKXL | 326 |
| Dd | | 447 | THLKTHSGRKMLIEIA-PTSNDP---GRNYCTAIRNRIGTFQOEYIELIADV----- | 495 |
| OY | | 327 | TCEVWGPFSPSKLMLEKLENKAQYKSKREP-----VWLNLNEAGWMOCCLSD | 374 |
| Dd | | 496 | -----PESPFGCVKIIEISQTTAKIS-FKPSPSHGCGPIHHAYQVDKEXVASETWKIVRSI | 548 |
| OY | | 375 | SGOVL-----LESNIKVLPTWSTPVHPRASALPAPPYG--SALPDQOTASAL-PDEPAAS | 426 |
| Dd | | 549 | GVOQTWVLSLEPN-----TTYEIRAAVANNGQGGQGYSKIETIQTLPVREPPSPSIH | 600 |
| OY | | 427 | ALPAA | 431 |
| Dd | | 601 | GQPSS | 605 |
| RESULT | | 24 | | |
| CEAS_HUMAN | | | | |
| ID | CEAS_HUMAN | STANDARD; | PRT; | 702 AA. |
| AC | P06731; | | | |
| DT | 01-JAN-1988 | (Rel. 06, Created) | | |
| DT | 01-DEC-1992 | (Rel. 24, Last sequence update) | | |
| DT | 15-MAR-2004 | (Rel. 43, Last annotation update) | | |
| DE | Carcinoembryonic antigen-related cell adhesion molecule 5 precursor | | | |
| DE | (Carcinoembryonic antigen) (CEA) (Meconium antigen I00) (CD66e | | | |
| DE | antigen). | | | |
| GN | CEACAM5 OR CEA. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RX | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=90258861; PubMed=2342461; | | | |
| RA | Schrewe H., Thompson J., Bona M., Hefta L.J.F., Maruya A., | | | |
| RA | Haseauer M., Shively J.E., von Kries S., Zimmermann W., | | | |
| RT | "Cloning of the complete gene for carcinoembryonic antigen: analysis | | | |
| RT | of its promoter indicates a region conveying cell type-specific | | | |
| RT | expression."; | | | |
| RL | Mol. Cell. Biol. 10:2738-2748(1990). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=86038676; PubMed=3670312; | | | |
| RA | Beauchemin N., Benchimol S., Courroyer D., Fuks A., Stannera C.P., | | | |
| RT | "Isolation and characterization of full-length functional cDNA clones | | | |
| RT | for human carcinoembryonic antigen." | | | |
| RL | Mol. Cell. Biol. 7:3221-3230(1987). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=89122014; PubMed=3220478; | | | |
| RA | Barnett T., Goebel S.J., Nothdurfft M.A., Elting J.J.; | | | |


```

RESULT 25
CAML_RAT STANDARD; PRT; 1259 AA.
ID CAML_RAT
AC 006595;
DT 01-FEB-1994 (Rel. 28, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neural cell adhesion molecule L1 precursor (N-CAM L1).
GN N1CAM OR CAML1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN NCBI_TaxId=10116;
RP [1]
SEQUENCE FROM N.A. (ISORFMS 1 AND 2).
RX MEDLINE=91372414; PUBMED=1894011;
RA Miura M., Kobayashi M., Asou H., Uyemura K.;
RT "Molecular cloning of cDNA encoding the rat neural cell adhesion
RT molecule L1. Two L1 isoforms in the cytoplasmic region are produced
RT by differential splicing.";
RL FEBS Lett. 289:91-95(1991).
CC -1- FUNCTION: Cell adhesion molecule with an important role in the
CC development of the nervous system. Involved in neuron-neuron
CC adhesion, neurite fasciculation, outgrowth of neurites, etc. Binds
CC to axonin on neurons.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q05695-1; Sequence=Displayed;
CC Name=2; Synonyms=Lics;
CC IsoId=Q05695-2; Sequence=VSP 002592;
CC -1- TISSUE SPECIFICITY: Isoform 2 is predominantly found in the brain,
CC while isoform 1 is found in the peripheral nervous system.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily.
CC L1/neurofascin/NeCAM family.
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 5 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X59149; CAA1860.1; -.
DR PIR, S36126; S36126.
DR HSP, P20241; 1CPB.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_C2.
DR Pfam; PF00041; fn3; 4.
DR Pfam; PF00047; Ig; 6.
DR SMART; SM00060; FN3; 4.
DR SMART; SM00408; IGc2; 5.
DR PROSITE; PS50835; IG_LIKE; 6.
DR Neurogeneis; Cell adhesion; Developmental protein; Glycoprotein;
DR Transmembrane; Repeat; Immunoglobulin domain; Signal;
KW Alternative splicing;
KW SIGNAL
FT CHAIN 1 1259
FT DOMAIN 20 1122
FT TRANSMEM 1123 1145
FT DOMAIN 1146 1259
FT DOMAIN 135 128
FT DOMAIN 138 225
FT DOMAIN 239 327
FT DOMAIN 332 419
FT DOMAIN 424 506
FT DOMAIN 517 600
IG-LIKE C2-TYPE 1.
IG-LIKE C2-TYPE 2.
IG-LIKE C2-TYPE 3.
IG-LIKE C2-TYPE 4.
IG-LIKE C2-TYPE 5.
IG-LIKE C2-TYPE 6.

```

```

FT DOMAIN 827 896
FT DOMAIN 932 994
FT DOMAIN 1032 1093
FT SITE 553 555
FT SITE 562 564
FT SITE 567 571
FT DISULFID 571 113
FT DISULFID 157 208
FT DISULFID 263 311
FT DISULFID 353 403
FT DISULFID 447 496
FT DISULFID 538 590
FT DISULFID 100 100
FT CARBOHYD 202 202
FT CARBOHYD 246 246
FT CARBOHYD 293 293
FT CARBOHYD 432 432
FT CARBOHYD 489 489
FT CARBOHYD 504 504
FT CARBOHYD 670 670
FT CARBOHYD 725 725
FT CARBOHYD 776 776
FT CARBOHYD 824 824
FT CARBOHYD 848 848
FT CARBOHYD 875 875
FT CARBOHYD 968 968
FT CARBOHYD 978 978
FT CARBOHYD 1021 1021
FT CARBOHYD 1029 1029
FT CARBOHYD 1072 1072
FT CARBOHYD 1106 1106
FT CARBOHYD 1179 1182
FT VARSPLIC 1179 1182
SQ SEQUENCE 1259 AA; 140934 MW; 0F12A7E4415F3C08 CR664;
Query Match 5.8%; Score 135; DB 1; Length 1259;
Best Local Similarity 20.2%; Pred. No. 0.19;
Matches 66; Conservative 58; Mismatches 114; Indels 88; Gaps 13;
QY 10 LLLVQLALPAAQGNKRVLGKGDYELTCTASQKKSIO-FHWKNSNOKILGNQGS 67
DB 507 ILAMIQVEATQIOGPRSTIEKKGARTFCQASFPDSIQASTWRDGR----- 557
QY 68 FLTGPSKLNDRASRSLMDQGNPLIKYKLIKEDSDTYC-----EVEDQKEVQLVLF 123
DB 558 -----DLQGRGSDKFTIEDGQ-LVINSLDVSDQDYSVCVASTLDEVESRAQLLV 608
QY 124 GLTAN-----SDPHLQGGSLTLTLESPGSSPVQCRSP-----RQKNIQGGKTL 169
DB 609 GSPGPVPHLESDHDLKQSGVHLSW-----SPAEDHNSPIETKYDIEFEDKEMAPKMF 662
QY 170 SVSQLELDQSGTWTCTYLQONQKVEFKIDIVL-----AFQKASSI 210
DB 663 SLGKV-----PQNQSTTLKLSPVYHYFRVLAINKYGGGSPVSEITVTPPEAAPKQPV 718
QY 211 VYKKGEGQ-----VERSP-LATFVE-KLTGSGELMMQARA-----SSK 249
DB 719 DVREGNETNMNMTWKPELRMDMNAPOIQRVQMRPLKQETWKEQTVSDPFLVSVTS 778
QY 250 SWITFDLKNKEVSVKRVYTDPKLQMG 275
DB 779 TFVPEYIKVQAVNNQGGKPEQVITIG 804
RESULT 26
SHS1_RAT STANDARD; PRT; 509 AA.
ID SHS1_RAT
AC P97710; 008951; 070426; 09QW15;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Protein-tyrosine phosphatase non-receptor type substrate 1 precursor
DE (SHP substrate-1) (SHPs-1) (Inhibitory receptor SHP-1) (Signal-
DE regulatory protein alpha-1) (Slrp-alpha-1) (Brain Ig-like molecule

```



```

FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 436 436 Y->F: ABOLISHES TYROSINE PHOSPHORYLATION AND PTEN11 BINDING; WHEN ASSOCIATED WITH F-460, F-477 AND F-501.
FT MUTAGEN 460 460 Y->F: ABOLISHES TYROSINE PHOSPHORYLATION AND PTEN11 BINDING; WHEN ASSOCIATED WITH F-436, F-477 AND F-501.
FT MUTAGEN 477 477 Y->F: STRONGLY REDUCES INSULIN-INDUCED TYROSINE PHOSPHORYLATION AND PTEN11 BINDING. ABOLISHES TYROSINE PHOSPHORYLATION AND PTEN11 BINDING; WHEN ASSOCIATED WITH F-436, F-477 AND F-501.
FT MUTAGEN 501 501 Y->F: STRONGLY REDUCES INSULIN-INDUCED TYROSINE PHOSPHORYLATION AND PTEN11 BINDING. ABOLISHES TYROSINE PHOSPHORYLATION AND PTEN11 BINDING; WHEN ASSOCIATED WITH F-436, F-477 AND F-501.
FT MUTAGEN 501 501 Y->F: STRONGLY REDUCES INSULIN-INDUCED TYROSINE PHOSPHORYLATION AND PTEN11 BINDING. ABOLISHES TYROSINE PHOSPHORYLATION AND PTEN11 BINDING; WHEN ASSOCIATED WITH F-436, F-477 AND F-501.
FT CONFLICT 8 8 MISSING (IN REF. 4).
FT CONFLICT 10 10 P -> L (IN REF. 4).
FT CONFLICT 25 25 F -> I (IN REF. 3).
FT CONFLICT 58 58 S -> C (IN REF. 4).
FT CONFLICT 99 100 KR -> MP (IN REF. 3; AA SEQUENCE).
FT CONFLICT 162 162 G -> A (IN REF. 2).
FT CONFLICT 189 189 D -> N (IN REF. 3).
FT CONFLICT 205 205 N -> L (IN REF. 3; AA SEQUENCE).
FT CONFLICT 209 209 N -> G (IN REF. 3; AA SEQUENCE).
FT CONFLICT 405 405 G -> F (IN REF. 3; AA SEQUENCE).
FT CONFLICT 416 416 E -> P (IN REF. 3; AA SEQUENCE).
FT CONFLICT 421 421 NARE -> EGON (IN REF. 3; AA SEQUENCE).
FT CONFLICT 438 438 R -> E (IN REF. 3; AA SEQUENCE).
FT CONFLICT 450 450 MISSING (IN REF. 3; AA SEQUENCE).
FT CONFLICT 499 499 MISSING (IN REF. 3; AA SEQUENCE).
SQ SEQUENCE 509 AA; 55690 MW; 581FEDAD429F4 CRC64;

Query Match 5.8%; Score 134.5; DB 1; Length 509;
Beet Local Similarity 22.0%; Pred. No. 0.062;
Matches 99; Conservative 75; Mismatches 150; Indels 127; Gaps 24;

```

```

DB 361 -----ETIPDNNAVMNFIQVACAL 364
RESULT 27
ID CAML_MOUSE STANDARD; PRT; 1260 AA.
AC P11627;
DT 01-OCT-1989 (Rel. 12, Created)
DE 15-MAR-2004 (Rel. 43, Last sequence update)
DE Neutral cell adhesion molecule l1 precursor (N-CAM L1).
OS L1CAM OR CAML.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC MEDLINE=88318924; PubMed=3412448;
RX TISSUE=Brain;
RA Moore M., Tacke R., Scherer H., Teplow D., Fruen K., Schachner M.;
RT "Neutral adhesion molecule l1 as a member of the immunoglobulin
RL superfamily with binding domains similar to fibronectin.";
CC Nature 334:701-703(1988).
CC -! FUNCTION: Cell adhesion molecule with an important role in the
CC development of the nervous system. Involved in neuron-neuron
CC adhesion, neurite fasciculation, outgrowth of neurites, etc. Binds
CC to axonin on neurons.
CC -! SUBCELLULAR LOCATION: Type I membrane protein.
CC -! SIMILARITY: Belongs to the immunoglobulin superfamily.
CC L1/neurofascin/NG2CAM family.
CC -! SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -! SIMILARITY: Contains 5 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X12875; CAA31368.1; -.
DR PIR; S05479; S05479.
DR HSP; P20241; ICFB.
DR MGD; MGI:96721; L1cam.
DR GO; GO:0007411; P:axon guidance; IMP.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR007110; Ig_1-like.
DR InterPro; IPR003598; Ig_C2.
DR Pfam; PF00041; fn3; 4.
DR SMART; PF00047; Ig; 6.
DR SMART; SMO0060; FN3; 2.
DR SMART; SMO0408; ICG2; 5.
DR PROSITE; PS50835; IG_LIKE; 6.
DR Neogenesis; Cell adhesion; Developmental protein; Glycoprotein;
DR Transmembrane; Repeat; Immunoglobulin domain; Signal.
DR SIGNAL 1 19
FT CHAIN 20 1260 NEURAL CELL ADHESION MOLECULE L1.
FT DOMAIN 20 1123 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1124 1146 POTENTIAL.
FT DOMAIN 1147 1260 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 138 225 IG-LIKE C2-TYPE 1.
FT DOMAIN 239 327 IG-LIKE C2-TYPE 2.
FT DOMAIN 332 419 IG-LIKE C2-TYPE 3.
FT DOMAIN 424 506 IG-LIKE C2-TYPE 4.
FT DOMAIN 517 600 IG-LIKE C2-TYPE 5.
FT DOMAIN 627 896 IG-LIKE C2-TYPE 6.
FT DOMAIN 932 994 FIBRONECTIN TYPE-III 1.
FT DOMAIN 1032 1094 FIBRONECTIN TYPE-III 2.
FT SITE 553 CELL ATTACHMENT SITE (POTENTIAL).

```

```

FT SITE 562 564 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 57 113 BY SIMILARITY.
FT DISULFID 157 208 BY SIMILARITY.
FT DISULFID 263 311 BY SIMILARITY.
FT DISULFID 353 403 BY SIMILARITY.
FT DISULFID 447 496 BY SIMILARITY.
FT DISULFID 538 590 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 725 725 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 824 824 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 848 848 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 875 875 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 978 978 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1030 1030 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1073 1073 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1107 1107 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1260 AA; 140968 MM; 22B57001CG2A538 CRG64;

```

```

Query Match 5.8%; Score 134.5; DB 1; Length 1260;
Best Local Similarity 20.1%; Pred. No. 0.2;
Matches 73; Conservative 59; Mismatches 140; Indels 91; Gaps 15;

```

```

QY 34 GDTVLTCTAQQKSIQFHMKNSQIKILNGSFLTGTGPKSKNDRAHSRLMDQGNFP 93
DB 346 GETARLDQVQVGRPOPEITWR-----INGSMETVYNKQKRYRI-EQGS-- 387
QY 94 LIINKLKIEDSDTYICEVEDOK-----EEVOLLVFGTLANSDFHL-LOGOSLTLTL 143
DB 388 LILSVQPTDTMNVNOCENRNGHLLANAYIVQQLPARILTKNGQTMVAVGSTRVYLIC 447
QY 144 ESPPSSSVQCRSRGKNI-----QGKTLVSQLELDQSGTWTCTVLQNKQKVEF 195
DB 448 KAFGAPVSVQWLDDEGTTVLQDERFFPYANGTILSRDLQANDGRYPCQANQNNVTI 507
QY 196 KIDIVYLAFOKASSI-----VYKKEGQVSESPFLATFVEKTLGSGSLMMQARASSK 249
DB 508 LANLQV---KEATQITGPRSAIEKKGARVTFTCQASFDPSL---QASITWRGDR--- 557
QY 250 SWITFDLKNKEVSVKRVTPQDPLQWKKLPHLTLPLQALPOVAGSGNLTLLAEAKTKLH 309
DB 558 -----DLQERGSDSKYFIEDGLVI-----QSL-DYSDQGNVSCVASFELDEVE 600
QY 310 QEVNLVWRA-----TQIQKNTCEVWGP-----TSPKMLSLKLENKEAK 350
DB 601 SRAQLLVVSGRPVPHLEISDRHLKQSVHLS---WSPADHNSPIEKYDIEFEDKEMA 657
QY 351 VSK 353
DB 658 PEK 660

```

```

RESULT 28
CAML_HUMAN
ID CAML_HUMAN STANDARD; PRT; 1257 AA.
AC P32004; Q8TA87;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neural cell adhesion molecule L1 precursor (N-CAM L1) (CD117 antigen).
GN L1CAM OR CAML1 OR NC15.
OS Homo sapiens (Human).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92031696; PubMed=1932117;
RA Kobayashi M., Mura M., Asou H., Uyemura K.;
RT "Molecular cloning of cell adhesion molecule L1 from human nervous
RT tissue: a comparison of the primary sequences of L1 molecules of
RT different origin.";
RL Biochim. Biophys. Acta 1090:238-240(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92120663; PubMed=1769655;
RA Hlavín M.L., Lemmon V.;
RT "Molecular structure and functional testing of human L1CAM: an
RT interspecies comparison.";
RL Genomics 11:416-423(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92329299; PubMed=1627459;
RA Reid R.A., Hemperly J.J.;
RT "Variants of human L1 cell adhesion molecule arise through alternate
RT splicing of RNA.";
RL J. Mol. Neurosci. 3:127-135(1992).
RN [4]
RP SEQUENCE FROM N.A.
RA Rosenthal A., Coutelle O., Drescher B.;
RL Submitted (Apr-1994) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=97432815; PubMed=9286695;
RA Brenner V., Nyakatura G., Rosenthal A., Platzter M.;
RT "Genomic organization of two novel genes on human Xq28: compact head
RT to head arrangement of IDH gamma and TRAP delta is conserved in rat
RT and mouse.";
RL Genomics 44:8-14(1997).
RN [6]
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RX MEDLINE=98147998; PubMed=9479034;
RA Coutelle O., Nyakatura G., Taudien S., Elgar G., Brenner S.,
RA Platzter M., Drescher B., Joutel M., Kenwright S., Rosenthal A.;
RT "The neural cell adhesion molecule L1: genomic organization and
RT differential splicing is conserved between man and the pufferfish
RT Fugu.";
RL Gene 208:7-15(1998).
RN [7]
RP SEQUENCE OF 20-36.
RX MEDLINE=88298876; PubMed=3136168;
RA Wolff J.M., Frank R., Mujoo K., Spiro R.C., Reisfeld R.A.,
RA Rathjen F.G.;
RT "A human brain glycoprotein related to the mouse cell adhesion
RT molecule L1.";
RL J. Biol. Chem. 263:11943-11947(1988).
RN [8]
RP SEQUENCE OF 332-371 FROM N.A.
RX MEDLINE=90353957; PubMed=2387585;
RA Djahali M., Matei M.-G., Nguyen C., Roux D., Demengeot J.,
RA Denizot F., Moos M., Schaefer M., Goridis C., Jordan B.R.;
RT "The gene encoding L1, a neural adhesion molecule of the
RT immunoglobulin family, is located on the X chromosome in mouse and
RT man.";
RL Genomics 7:587-593(1990).
RN [9]
RP SEQUENCE OF 353-1176 FROM N.A.
RC TISSUE=fetal brain;
RX MEDLINE=92020233; PubMed=1923824;
RA Rosenthal A., Mackinnon R.N., Jones D.S.C.;
RT "PCR walking from microdissection clone NS4 identifies three exons
RT from the human gene for the neural cell adhesion molecule L1
RT (CAM-L1).";
RL Nucleic Acids Res. 19:5395-5401(1991).

```

[10] SEQUENCE OF 809-1257 FROM N.A.
RA TISSUE=Pancreas;
RC MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Dieckmann L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepiencko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywnicki M.I., Skalska T., Smallie D.E.,
RA Smerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
[11] SEQUENCE OF 1030-1257 FROM N.A.
RA MEDLINE=91132183; PubMed=1993895;
RX Harper J.R., Prince J.T., Healy P.A., Stuart J.K., Nauman S.J.,
RA Stumpel W.B.,
RT "Isolation and sequence of partial cDNA clones of human L1: homology
RL of human and rodent L1 in the cytoplasmic region.";
RN J. Neurochem. 56:797-804(1991).
[12] VARIANT HSAS TYR-264.
RX MEDLINE=94004956; PubMed=8401576;
RA Jouet M., Roenthal A., Macfarlane J., Kenrick S., Domai D.,
RT "A missense mutation confirms the L1 defect in X-linked hydrocephalus
RL (HSAS).";
RN Nat. Genet. 4:331-331(1993).
[13] VARIANT HSAS/MASA LEU-1194.
RX MEDLINE=95187172; PubMed=788131;
RA Fransén E., Schrander-Stumpel C., Vils L., Coucke P., van Camp G.,
RA Williams P.J.,
RT "X-linked hydrocephalus and MASA syndrome present in one family are
RL due to a single missense mutation in exon 28 of the L1CAM gene.";
RN Hum. Mol. Genet. 3:2255-2256(1994).
[14] VARIANT HSAS GLN-184 AND ARG-452, AND VARIANT MASA GLN-210.
RX MEDLINE=95004608; PubMed=7920659;
RA Jouet M., Roenthal A., Armstrong G., Macfarlane J., Stevenson R.,
RA Peterson J., Metzberg A.A., Ionescu V., Temple K., Kenrick S.,
RT "X-linked spastic paraplegia (SPPL1), MASA syndrome and X-linked
RL hydrocephalus result from mutations in the L1 gene.";
RN Nat. Genet. 7:402-407(1994).
[15] VARIANTS MASA GLN-210 AND ASN-598.
RX MEDLINE=95004609; PubMed=7920660;
RA Vils L., van Camp G., Coucke P., Fransén E., de Boule K.,
RA Reyniers E., Korn B., Poueika A., Wilson G., Schrander-Stumpel C.,
RA Winter R.M., Schwartz C., Willems P.J.,
RT "MASA syndrome is due to mutations in the neural cell adhesion gene
RL L1CAM.";
RN Nat. Genet. 7:408-413(1994).
[16] VARIANTS HSAS/MASA SER-9; SER-121; LYS-309; PHE-768; LEU-941 AND
RN CYS-1070.
RX MEDLINE=95282776; PubMed=7762552;
RA Jouet M., Moncla A., Peterson J., McKernan C., Fryer A., Carpenter N.,
RA Holmberg E., Madlous C., Kenrick S.,
RT "New domains of neural cell-adhesion molecule L1 implicated in
RL X-linked hydrocephalus and MASA syndrome.";
RN Am. J. Hum. Genet. 56:1304-1314(1995).

[17]
RP VARIANTS HSAS/MASA GLN-184, GLN-210, TYR-264; ARG-452, ASN-598 AND LEU-1194.
RX MEDLINE=96153146; PubMed=8556302;
RA F. Fransen E., Lemmon V., van Camp G., Vits L., Coucke P., Willems P.J.;
RT "CRASH syndrome: clinical spectrum of corpus callosum hypoplasia,
RT retardation, adducted thumbs, spastic paraparesis and hydrocephalus
RT due to mutations in one single gene, LI-1";
RN Eur. J. Hum. Genet. 3:273-284(1995).
[18]
RN ERRATUM.
RA Fransen E., Lemmon V., van Camp G., Vits L., Coucke P., Willems P.J.;
RL Eur. J. Hum. Genet. 4:126-126(1996).
[19]
RP VARIANTS HSAS/MASA/SPG1 SER-179 AND ARG-370.
RX MEDLINE=96057511, PubMed=7562969;
RA Ruiz J.C., Cuppens H., Legius E., Fryns J.-P., Glover T., Marynen P.,
RA Cassiman J.-J.;
RT "Mutations in LI-CAM in two families with X linked complicated
RT spastic paraplegia, MASH syndrome, and HSAS.";
RL J. Med. Genet. 32:549-552(1995).
[20]
RP VARIANTS HSAS CYS-194 AND LEU-240.
RX MEDLINE=97083370, PubMed=8929944;
RA Gu S.-H., Orich U., Veske A., Enders H., Kluender K., Schloesser M.,
RA Engel W., Schwinger E., Gal A.;
RT "Five novel mutations in the LICAM gene in families with X linked
RT hydrocephalus.";
RL J. Med. Genet. 33:103-106(1996).
[21]
RN VARIANTS HSAS GLN-184, 439-VAL-THR-443 DEL; CYS-784 AND
RP 936-LEU-LEU-948 DEL.
RX MEDLINE=97338664, PubMed=9195224;
RA Macfarlane J.R., Du J.-S., Peppy M.E., Ramsden S., Donna D.,
RA Charlton R., Garrett C., Tolmie J., Yates J.R.W., Berry C., Goudie D.,
RA Monica A., Hunt P., Hodgson S., Joutet M., Kenwick S.;
RT "Nine novel LI CAM mutations in families with X-linked
RT hydrocephalus.";
RL Hum. Mutat. 9:512-518(1997).
[22]
RP VARIANTS HSAS/ASA ASP-691; ARG-698 AND PRO-935.
RX MEDLINE=96180721; PubMed=9521242;
RA Du Y.-L., Srivastava A.K., Schwartz C.E.;
RT "Multiple exon screening using restriction endonuclease
RT fingerprinting (REF): detection of six novel mutations in the LI cell
RT adhesion molecule (LICAM) gene.";

```

Db 687 FRVTAINKYGEPSVSEVTVPTEAEAKNPVDKGEINET-----TNWVI 733
Oy 317 MRATOLQKNTCEVWGP-----TSPKMLSLKLENKEAKVSKKEKPYWVLPNPAQMWQ- 369
Db 734 -----TWKPLRMNDNAPQVQYRVQ-----WRQGTGRGKQOE 765
Oy 370 CLSDSGVLLSNIKVLPTWSTVPVPRASLPAPPTGSALPDQTSALPDPPASALP 429
Db 766 QIVSD--PFLVVSNTSTP-----VPYEIKVQAVNSQKG---PEQVTVTGYSGEDYPPAIP 816

RESULT 29
KMLS_CHICK STANDARD: PRT, 1906 AA.
ID KMLS_CHICK
AC P11799; P19038;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myosin light chain kinase, smooth muscle and non-muscle isozymes
DE (EC 2.7.1.117) (MCK) [Contains: Telokin].
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM MCK-210).
RX MEDLINE=96033976; PubMed=7589469;
RA Waterson D.M., Collinge M., Lukas T.J., van Eldik L.J., Birukov K.G.,
RA Stepanova O.V., Shirinsky V.P.;
RT "Multiple gene products are produced from a novel protein kinase
RT transacription region.";
RT FEBS Lett. 373:217-220(1995).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM MCK-108).
RX MEDLINE=90192792; PubMed=231520;
RA Olson N.J., Pearson R.B., Needleman D.S., Hurwitz M.J., Kemp B.E.,
RA Means A.R.;
RT "Regulatory and structural motifs of chicken gizzard myosin light
RT chain kinase.";
RT Proc. Natl. Acad. Sci. U.S.A. 87:2284-2288(1990).
RN [3]
RP SEQUENCE OF 649-1906 FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Fibroblast;
RX MEDLINE=90361738; PubMed=2202734;
RA Shoemaker M.O., Lau W., Shatuck R.L., Kwiatkowski A.P.,
RA Marietian P.E., Guerra-Santos L., Wilson E., Lukas T.J.,
RA van Eldik L.J., Waterson D.M.;
RT "Use of DNA sequence and mutant analyses and antisense
RT oligodeoxynucleotides to examine the molecular basis of nonmuscle
RT myosin light chain kinase autoinhibition, calmodulin recognition, and
RT activity.";
RT J. Cell Biol. 111:1107-1125(1990).
RN [4]
RP SEQUENCE OF 1259-1906 FROM N.A.
RC TISSUE=Gizzard;
RX MEDLINE=87157587; PubMed=3030394;
RA Guerriero V. Jr., Russo M.A., Olson N.J., Putkey J.A., Means A.R.;
RT "Domain organization of chicken gizzard myosin light chain kinase
RT deduced from a cloned cDNA.";
RT Biochemistry 25:8372-8381(1986).
RN [5]
RP SEQUENCE OF 1750-1906 FROM N.A. (ISOFORM TELOKIN).
RC TISSUE=Gizzard;
RX MEDLINE=93073972; PubMed=1444462;
RA Yoshikai S., Ikebe W.;
RT "Molecular cloning of the chicken gizzard telokin gene and cDNA.";
RT Arch. Biochem. Biophys. 299:242-247(1992).
RN [6]
RP SEQUENCE OF 1750-1906 FROM N.A. (ISOFORM TELOKIN).
RX MEDLINE=92236611; PubMed=1373815;
RA Collinge M., Marietian P.E., Zimmer W.E., Shattuck R.L., Lukas T.J.,
RA van Eldik L.J., Waterson D.M.;

```

```

RT "structure and expression of a calcium-binding protein gene contained
RT within a calmodulin-regulated protein kinase gene.";
RL Mol. Cell. Biol. 12:2359-2371(1992).
CC -1- FUNCTION: Phosphorylates a specific serine in the N-terminus of a
CC myosin light chain, which leads to the formation calmodulin/MCK
CC signal transduction complexes which allow selective transduction
CC of calcium signals.
CC -1- CATALYTIC ACTIVITY: ATP + [myosin light-chain] = ADP + [myosin
CC light-chain] phosphate.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative Initiation;
CC MCK-108/Smooth-muscle and Telokin, are produced by alternative
CC Initiation;
CC -1- TISSUE SPECIFICITY: ISOFORM TELOKIN IS EXPRESSED IN GIZZARD,
CC HEART, LUNG, INTESTINE, AND SKELETAL MUSCLE ALTHOUGH THE LEVELS OF
CC THE EXPRESSION IN THE LATTER WERE MUCH LESS THAN THAT IN THE
CC GIZZARD.
CC -1- DOMAIN: TELOKIN BINDS CALMODULIN.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -1- SIMILARITY: Contains 9 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 fibronectin type III domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52876; CA37056.1; -
DR EMBL; X52876; CA37057.1; -
DR EMBL; X52876; CA37058.1; -
DR EMBL; M31048; AAA49069.1; -
DR EMBL; M14953; AAA69964.1; -
DR EMBL; M96655; AAA49083.1; -
DR EMBL; M88283; AAA48647.1; -
DR EMBL; M88284; AAB53768.1; -
DR PIR; S68235; S68235.
DR PDB; 1CDL; 31-ANG-94.
DR PDB; 1VRK; 27-APR-99.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR007110; IG_1-like.
DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR007119; PROT_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00047; fn3; 1.
DR Pfam; PF00047; ig; 9.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00408; IGC2; 8.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PSS0835; IG_LIKE; 9.
DR PROSITE; PSS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR Transferrase; Serine/threonine-protein kinase; Calmodulin-binding;
KW ATP-binding; Phosphorylation; Immunoglobulin domain; Repeat;
KW Alternative initiation; 3D-structure.
FT CHAIN 1 1906
FT CHAIN 935 1906
FT CHAIN 1750 1906
FT CHAIN 1750 1906
FT INT_MET 935 935
FT INT_MET 1750 1750
FT DOMAIN 28 117
FT DOMAIN 156 244
FT DOMAIN 156 244

```

```

FT DOMAIN 429 517 IG-LIKE C2-TYPE 3.
FT DOMAIN 521 613 IG-LIKE C2-TYPE 4.
FT DOMAIN 637 725 IG-LIKE C2-TYPE 5.
FT DOMAIN 735 830 IG-LIKE C2-TYPE 6.
FT DOMAIN 1084 1172 IG-LIKE C2-TYPE 7.
FT DOMAIN 1225 1313 IG-LIKE C2-TYPE 8.
FT DOMAIN 1330 1400 FIBRONECTIN TYPE-III.
FT DOMAIN 1453 1708 PROTEIN KINASE.
FT DOMAIN 1794 1885 IG-LIKE C2-TYPE 9.
FT NP_BIND 1459 1467 ATP (BY SIMILARITY).
FT BINDING 1462 1462 ATP (BY SIMILARITY).
FT ACT_SITE 1574 1574 CALMODULIN AUTOINHIBITION (AM13) REGION
FT DOMAIN 1716 1728 (POTENTIAL).
FT FT CALMODULIN RECOGNITION (RS20) REGION
FT DOMAIN 1730 1749 (POTENTIAL).
FT FT MOTIF IA.
FT DOMAIN 1317 1364 MOTIF IB.
FT DOMAIN 1385 1402 4 X REPEATS, MOTIF IIA.
FT DOMAIN 660 1833 IIA-1.
FT REPEAT 660 676 IIA-2.
FT REPEAT 758 774 IIA-3.
FT REPEAT 1107 1123 IIA-4.
FT REPEAT 1817 1833 IIA-4.
FT DOMAIN 693 1866 5 X REPEATS, MOTIF IIB.
FT REPEAT 693 708 IIB-1.
FT REPEAT 791 807 IIB-2.
FT REPEAT 1140 1156 IIB-3.
FT REPEAT 1281 1287 IIB-4.
FT REPEAT 1851 1866 IIB-5.
FT DOMAIN 970 1226 4 X REPEATS, MOTIF III.
FT REPEAT 970 987 III-1.
FT REPEAT 999 1016 III-2.
FT REPEAT 1061 1078 III-3.
FT REPEAT 1209 1226 III-4.
FT DOMAIN 1700 1763 CALMODULIN-BINDING.
FT DOMAIN 1896 1906 POLY-GLU.
FT MOD_RES 1748 1748 PHOSPHORYLATION.
FT MOD_RES 1762 1762 PHOSPHORYLATION.
FT CONFLICT 1439 1439 R -> O (IN REP. 4)
SQ SEQUENCE 1906 AA; 210445 MW; AD7DBA3869E33363 CRC64;

Query Match 5.6%; Score 130; DB 1; Length 1906;
Best Local Similarity 20.8%; Pred. No. 0.7;
Matches 82; Conservative 58; Mismatches 142; Indels 112; Gaps 15;

QY 89 OGNFLLIKNLKIENSDTYICEVD---QKEEVLLVFGLTNSDTHLQCGSLTTLLE 144
DB 81 RGFSLVTKGVQEGDSKTYTCEANDGVRQVTVETVEG---NS---LKKYSLPSSAK 133
QY 145 SPPG---SSPSVQCR-----SPRGKNIQGGKT----- 168
DB 134 TPQGLSVPEVHRPSINGESPSPKATKPNRVVBEQGTGRSPSKITGRRPOQVTWTKGD 193
QY 169 -----LSVSOLELQDSGTWTCTVLQNKKEVFKIDIV----- 201
DB 194 IHLQGNERNRMFEKIGIYLEIQNVQLADAGIYITCTVNSAKKASVSAELTVQGRDKDT 253
QY 202 -----LA-----FOKASSIVYKKEGEQVSEFPLAFVTEKUTSGGELMW 240
DB 254 HAOPLCMPKPTTLTKALENSDFQKATSNGLAKELKSTLSTEL-VWETKDRLSAKKETFFY 312
QY 241 OABERASSSSKSWITTPFLKNKEVSVKRVTDPKLQMKKULPLHLTLPOALPOVAGSGNLTLA 300
DB 313 TSREAKDQKOGNOENANVPLQESRGTKGPOVLQKTSSTTLQAVVAQDEPPAEPOTTFI 372
QY 301 LEAKTKLHQEVNVLVVMRATQLOKNUJCEVMGPTSPKMLSLKLEN-----KEAKVSKRE 355
DB 373 RQAEARK--RTVQPLMTTITTOENPSLT-----QGVSPR---SRETFENRAGVRKSVKEKRE 423
QY 356 KQVWVLNPNAGMWQCLSDSGO-VLLESNIKLP 388
DB 424 -PLGIPIQFESRPSQSLSESGOEIKFKSVKSGKP 456

```

```

RESULT 30
PGEM_MOUSE
ID PGEM_MOUSE STANDARD; PRT; 3707 AA.
AC 005793;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Basement membrane-specific heparan sulfate proteoglycan core
DE protein precursor (HSPG) (perlecan) (PLC).
GN HSPG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=92078153; Pubmed=1744087;
RA Noonan D.M., Fulle A., Valente P., Cai S., Horigan E., Sasaki M.,
RA Yamada Y., Haasell J.R.,
RT "The complete sequence of perlecan, a basement membrane heparan
RT sulfate proteoglycan, reveals extensive similarity with laminin A
RT chain, low density lipoprotein-receptor, and the neural cell adhesion
RT molecule."
RL J. Biol. Chem. 266:22939-22947(1991).
RN [2]
RP SEQUENCE OF 940-1601 AND 1870-2600 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=89034110; Pubmed=2972708;
RA Noonan D.M., Horigan E.A., Ledbetter S.R., Vogel G., Sasaki M.,
RA Yamada Y., Haasell J.R.;
RT "Identification of cDNA clones encoding different domains of the
RT basement membrane heparan sulfate proteoglycan."
RL J. Biol. Chem. 263:16379-16387(1988).
CC -1- FUNCTION: This protein is an integral component of basement
CC membranes. It is responsible for the fixed negative electrostatic
CC charge and is involved in the charge-selective ultrafiltration
CC properties. It serves as an attachment substrate for cells.
CC -1- SUBUNIT: Purified perlecan has a strong tendency to aggregate in
CC dimers or stellate structures. It interacts with other basement
CC membrane components such as laminin, prolargin and collagen type
CC IV.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: Found in the basement membranes.
CC -1- PTRN: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
CC AND O-LINKED OLIGOSACCHARIDES.
CC -1- SIMILARITY: Contains 4 LDP-receptor class A domains.
CC -1- SIMILARITY: Contains 11 laminin EGF-like domains.
CC -1- SIMILARITY: Contains 3 laminin IV domains.
CC -1- SIMILARITY: Contains 15 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 3 laminin G-like domains.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 SEA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M77174; AAA39911.1; -
DR EMBL; J04054; AAA39899.1; -
DR EMBL; J04055; AAA39912.1; -
DR PIR; S18252; S18252.
DR PDB; 1GL4; 28-NOV-01.
DR MGD; MGI:96257; Hspg2.
DR GO; GO:0005604; C:basement membrane; IDA.
DR GO; GO:0008104; P:protein localization; IMP.
DR InterPro; IPR008985; Cona_1like_1ec_5l.
DR InterPro; IPR00742; EGF_2.
DR InterPro; IPR006209; EGF_1like.

```

DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000034; Laminin_B.
 DR InterPro; IPR002049; Laminin_EGF.
 DR InterPro; IPR001791; Laminin_G.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR InterPro; IPR000082; SEA_domain.
 DR Pfam; PF000008; EGF_4.
 DR Pfam; PF000047; Ig; 15.
 DR Pfam; PF00052; Laminin_B; 3.
 DR Pfam; PF00053; Laminin_EGF; 7.
 DR Pfam; PF00054; Laminin_G; 3.
 DR Pfam; PF00057; LDL_recept_a; 4.
 DR Pfam; PF01390; SEA; 1.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR ProDom; PD001031; Laminin_B; 3.
 DR SMART; SM00180; EGF_Lam; 7.
 DR SMART; SM00408; IGC2; 14.
 DR SMART; SM00281; LamB; 3.
 DR SMART; SM00282; LamG; 3.
 DR SMART; SM00192; LDLa; 4.
 DR SMART; SM00200; SEA; 1.
 DR PROSITE; PS00022; EGF_1; 8.
 DR PROSITE; PS01186; EGF_2; 5.
 DR PROSITE; PS50026; EGF_3; 4.
 DR PROSITE; PS50835; IG_LIKE; 15.
 DR PROSITE; PS50025; LAM_G_DOMAIN; 3.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
 DR PROSITE; PS01209; LDLRA_1; 4.
 DR PROSITE; PS50068; LDLRA_2; 4.
 DR PROSITE; PS50024; SEA; 1.
 DR Signal; Basement membrane; Proteoglycan; Repeat; Glycoprotein;
 KW Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain;
 KW Extracellular matrix; EGF-like domain; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 3707
 FT DOMAIN 80 194
 FT DOMAIN 195 234
 FT DOMAIN 281 319
 FT DOMAIN 320 359
 FT DOMAIN 360 403
 FT DOMAIN 404 504
 FT DOMAIN 521 530
 FT DOMAIN 531 730
 FT DOMAIN 731 763
 FT DOMAIN 764 813
 FT DOMAIN 814 871
 FT DOMAIN 879 923
 FT DOMAIN 924 933
 FT DOMAIN 934 1125
 FT DOMAIN 1126 1158
 FT DOMAIN 1159 1208
 FT DOMAIN 1209 1265
 FT DOMAIN 1275 1324
 FT DOMAIN 1325 1334
 FT DOMAIN 1335 1529
 FT DOMAIN 1530 1562
 FT DOMAIN 1563 1612
 FT DOMAIN 1613 1670
 FT DOMAIN 1671 1771
 FT DOMAIN 1772 1865
 FT DOMAIN 1866 1954
 FT DOMAIN 1955 2049
 FT DOMAIN 2050 2148
 FT DOMAIN 2149 2244
 FT DOMAIN 2245 2343
 FT DOMAIN 2344 2436
 FT DOMAIN 2437 2532
 FT DOMAIN 2533 2619
 FT DOMAIN 2620 2720
 FT DOMAIN 2721 2809
 FT DOMAIN 2810 2895
 BASEMENT MEMBRANE-SPECIFIC HEPARAN
 SULFATE PROTEOGLYCAN CORE PROTEIN.
 SEA.
 LDL-RECEPTOR CLASS A 1.
 LDL-RECEPTOR CLASS A 2.
 LDL-RECEPTOR CLASS A 3.
 LDL-RECEPTOR CLASS A 4.
 IG-LIKE C2-TYPE 1.
 LAMININ EGF-LIKE 1 (N-TERMINAL).
 LAMININ EGF-LIKE 1 (DOMAIN III A).
 LAMININ EGF-LIKE 1 (C-TERMINAL).
 LAMININ EGF-LIKE 2.
 LAMININ EGF-LIKE 3 (INCOMPLETE).
 LAMININ EGF-LIKE 5 (N-TERMINAL).
 LAMININ EGF-LIKE 5 (DOMAIN III B).
 LAMININ EGF-LIKE 5 (C-TERMINAL).
 LAMININ EGF-LIKE 6.
 LAMININ EGF-LIKE 7.
 LAMININ EGF-LIKE 8.
 LAMININ EGF-LIKE 9 (N-TERMINAL).
 LAMININ EGF-LIKE 9 (DOMAIN III C).
 LAMININ EGF-LIKE 9 (C-TERMINAL).
 LAMININ EGF-LIKE 10.
 LAMININ EGF-LIKE 11.
 IG-LIKE C2-TYPE 2.
 IG-LIKE C2-TYPE 3.
 IG-LIKE C2-TYPE 4.
 IG-LIKE C2-TYPE 5.
 IG-LIKE C2-TYPE 6.
 IG-LIKE C2-TYPE 7.
 IG-LIKE C2-TYPE 8.
 IG-LIKE C2-TYPE 9.
 IG-LIKE C2-TYPE 10.
 IG-LIKE C2-TYPE 11.
 IG-LIKE C2-TYPE 12.
 IG-LIKE C2-TYPE 13.
 IG-LIKE C2-TYPE 14.

FT DOMAIN 2896 2980 IG-LIKE C2-TYPE 15.
 FT DOMAIN 2984 3162 LAMININ G-LIKE 1.
 FT DOMAIN 3163 3241 EGF-LIKE.
 FT DOMAIN 3245 3425 LAMININ G-LIKE 2.
 FT DOMAIN 3518 3705 LAMININ G-LIKE 3.
 FT SITE 65 67 HEPARAN SULFATE (POTENTIAL).
 FT SITE 71 73 HEPARAN SULFATE (POTENTIAL).
 FT SITE 76 78 HEPARAN SULFATE (POTENTIAL).
 FT SITE 3615 3617 MEDIATES MOTOR NEURON ATTACHMENT (POTENTIAL).
 FT DISULFID 199 212 BY SIMILARITY.
 FT DISULFID 206 225 BY SIMILARITY.
 FT DISULFID 219 234 BY SIMILARITY.
 FT DISULFID 285 297 BY SIMILARITY.
 FT DISULFID 292 310 BY SIMILARITY.
 FT DISULFID 304 319 BY SIMILARITY.
 FT DISULFID 325 337 BY SIMILARITY.
 FT DISULFID 332 350 BY SIMILARITY.
 FT DISULFID 344 359 BY SIMILARITY.
 FT DISULFID 368 381 BY SIMILARITY.
 FT DISULFID 375 394 BY SIMILARITY.
 FT DISULFID 388 403 BY SIMILARITY.
 FT DISULFID 428 479 BY SIMILARITY.
 FT DISULFID 764 773 BY SIMILARITY.
 FT DISULFID 766 780 BY SIMILARITY.
 FT DISULFID 783 792 BY SIMILARITY.
 FT DISULFID 795 811 BY SIMILARITY.
 FT DISULFID 814 829 BY SIMILARITY.
 FT DISULFID 816 839 BY SIMILARITY.
 FT DISULFID 842 851 BY SIMILARITY.
 FT DISULFID 854 869 BY SIMILARITY.
 FT DISULFID 1159 1168 BY SIMILARITY.
 FT DISULFID 1161 1175 BY SIMILARITY.
 FT DISULFID 1178 1187 BY SIMILARITY.
 FT DISULFID 1190 1206 BY SIMILARITY.
 FT DISULFID 1209 1224 BY SIMILARITY.
 FT DISULFID 1211 1234 BY SIMILARITY.
 FT DISULFID 1237 1246 BY SIMILARITY.
 FT DISULFID 1249 1263 BY SIMILARITY.
 FT DISULFID 1275 1287 BY SIMILARITY.
 FT DISULFID 1277 1293 BY SIMILARITY.
 FT DISULFID 1295 1304 BY SIMILARITY.
 FT DISULFID 1307 1322 BY SIMILARITY.
 FT DISULFID 1563 1572 BY SIMILARITY.
 FT DISULFID 1565 1579 BY SIMILARITY.
 FT DISULFID 1582 1591 BY SIMILARITY.
 FT DISULFID 1594 1610 BY SIMILARITY.
 FT DISULFID 1613 1628 BY SIMILARITY.
 FT DISULFID 1615 1638 BY SIMILARITY.
 FT DISULFID 1641 1650 BY SIMILARITY.
 FT DISULFID 1653 1668 BY SIMILARITY.
 FT DISULFID 1792 1839 BY SIMILARITY.
 FT DISULFID 1886 1932 BY SIMILARITY.
 FT DISULFID 1976 2021 BY SIMILARITY.
 FT DISULFID 2073 2118 BY SIMILARITY.
 FT DISULFID 2170 2215 BY SIMILARITY.
 FT DISULFID 2268 2313 BY SIMILARITY.
 Query Match 5.5%; Score 127; DB 1; Length 3707;
 Best Local Similarity 19.3%; Pred. No. 2; Indels 168; Gaps 22;
 Matches 100; Conservative 66; Mismatches 185;

QY 32 KKGDVLELTCTASOKSIQFMKNSNGIKILNGSGFLTGKPSKLANDRAISRSLMDQGN 91
 Db 2447 KMGDITLBECLSSGSEPRSSPRWTR-----LGI-----PVKLEPRMFGLAN-----SH 2488
 QY 92 PPLIKNLKIEDSDPYICEVED---OKEEVOLLVFGLTANSDDTLHOGQSLVLTLSPP 147
 Db 2489 AMKLIAVKRSDAGTYVCGAQNALGTAKQVELVDGTGVAAPGTPOVQVESESLTLEA-- 2546
 QY 148 GSSPSVOCRS-----PRGKNIOGKTLISVSQLELDSDSGTWCTCVLONQ 190
 Db 2547 GHTATLHCATGNGPPTIHWSKLRAPLPMOHRIE-GMTLVIPVAODSDSGOYICNATNSA 2605

| | | | | | | |
|----|------|----------------|--------------|--------------------|-----------------|-------|
| QY | 191 | KKVEFKIDIVLVA | PKQKASSI--- | YKKKEGEQVEFSF----- | PLAF----- | 227 |
| Db | 2606 | GHTETAVVLHES | PEPYATIIIEPH | TSAOEGNVLQLOCLAH | GTPLPYQMSLVGVLP | 2665 |
| QY | 228 | ----- | ----- | ----- | ----- | ----- |
| Db | 2666 | VRNQLRLREPT | VPEDSGRYRCOV | SNRVSAAEFAQVL | VOGSSSNLPDISIG | 2725 |
| QY | 267 | TQDPKIQ--- | MGKULPLHLTL | PQA---L | PQYAGSGNLTAL | 319 |
| Db | 2726 | T--TQLEFRNIG | ASVEFFCAVBN | ERGTHLRMLKEG | QGLPQGHVSVDG-- | 2774 |
| QY | 320 | TQLOKN----- | LTCEWGP----- | ----- | ----- | 351 |
| Db | 2775 | ONLQONCOGTY | YVCOAHMGQAQ | ATQALIVQALP | SVLIVNTSVHSV | 2834 |
| QY | 352 | -SKREKPYVW---- | LNP----- | ----- | ----- | ----- |
| Db | 2835 | GDPRPQVYWS | KVGGILRGIV | QSGIILIAVEL | ADAQYCAATNA | 2925 |
| QY | 384 | IKVLEPTWST | PVHPRASALP | ADPTGSALP | DPQTASALPDP | 422 |
| Db | 2893 | VQALPQISTP | PEIRY----- | PAGSAVFP | CMASGYPRP | 2925 |

RESULT 31

| | | | | |
|----|--|-----------|------|---------|
| ID | NCAM2_MOUSE | STANDARD; | PRT; | 725 AA. |
| AC | P13594; Q61950; | | | |
| DT | 01-JAN-1990 (Rel. 13, created) | | | |
| DT | 01-JAN-1990 (Rel. 13, last sequence update) | | | |
| DT | 10-OCT-2003 (Rel. 42, last annotation update) | | | |
| DE | Neural cell adhesion molecule 1, 120 kDa isoform precursor (N-CAM 120) | | | |
| DE | (NCAM-120). | | | |
| GN | NCAM1 OR NCAM. | | | |
| OS | Mus musculus (Mouse). | | | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. | | | |
| OX | NCBI_TaxId=10090; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=C57BL/6; | | | |
| RX | MEDLINE=87246524; PubMed=3595563; | | | |
| RA | Barthele D., Santoni M.-J., Wille M., Ruppert C., Calx J.-C., | | | |
| RA | Hirsch M.-R., Fontecilla-Camps J.-C., Goridis C.; | | | |
| RT | "Isolation and nucleotide sequence of mouse NCAM cDNA that codes for | | | |
| RL | a Mr 79,000 polypeptide without a membrane-spanning region."; | | | |
| RN | EMBO J. 6:907-914(1987). | | | |
| RN | [2] | | | |
| RP | SEQUENCE OF 20-700 FROM N.A. | | | |
| RC | STRAIN=C57BL/6; TISSUE=Brain; | | | |
| RX | MEDLINE=89251563; PubMed=2721486; | | | |
| RA | Santoni M.-J., Barthele D., Vopper G., Boned A., Goridis C., Wille M.; | | | |
| RT | "Differential exon usage involving an unusual splicing mechanism | | | |
| RL | generated at least eight types of NCAM cDNA in mouse brain."; | | | |
| RN | EMBO J. 8:385-392(1989). | | | |
| RN | [3] | | | |
| RP | SEQUENCE OF 642-725 FROM N.A. | | | |
| RX | MEDLINE=86283628; PubMed=3396534; | | | |
| RA | Barbais J.A., Chaix J.C., Steimeyer M., Goridis C.; | | | |
| RT | "Differential splicing and alternative polyadenylation generates | | | |
| RL | distinct NCAM transcripts and proteins in the mouse."; | | | |
| RN | EMBO J. 7:625-632(1988). | | | |
| RN | [4] | | | |
| RP | SEQUENCE OF 20-36. | | | |
| RX | MEDLINE=86140120; PubMed=3512556; | | | |
| RA | Rougon G., Marshak D.R.; | | | |
| RT | "Structural and immunological characterization of the amino-terminal | | | |
| RL | domain of mammalian neural cell adhesion molecules."; | | | |
| RN | J. Biol. Chem. 261:3396-3401(1986). | | | |
| CC | -1- FUNCTION: This protein is a cell adhesion molecule involved in | | | |
| CC | neuron-neuron adhesion, neurite fasciculation, outgrowth of | | | |

CC neites, etc.
CC -I- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC
CC -I- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=N-CAM 120;
CC IsoId=P13594-1; Sequence=Displayed;
CC Name=N-CAM 180;
CC IsoId=P13595-1; Sequence=External;
CC Name=N-CAM 140;
CC IsoId=P13595-2; Sequence=External!
CC -I- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -I- SIMILARITY: Contains 2 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its contents is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isdb.ch/announce/
CC or send an email to license@isdb.ch).
CC -----
CD EMBL; Y00051; CA668263.1; -.
DR EMBL; X15049; CA31348.1; ALT_SEQ.
DR EMBL; X07195; CA31073.1; -.
PIR; A29673; IUMSNG.
DR PDB; 2NCM; 12-MAR-97.
DR PDB; 3NCM; 23-JUL-99.
MGD; MGI:97281; Ncaml.
DR InterPro; IPRO08957; FN_III-like.
DR InterPro; IPRO03961; FN_III.
DR InterPro; IPRO07110; Ig_Ilike.
DR InterPro; IPRO03598; Ig_C2.
Pfam; PF00041; fn3; 2.
PFam; PF00047; Ig; 5.
SMART; SMO0408; FN3; 2.
SMART; SMO0408; IGc2; 5.
PROSITE; PS50835; IG_LIKE; 5.
Kw Cell adhesion; Glycoprotein; Repeat; Alternative splicing;
Km Immunoglobulin domain; Signal; Heparin-binding; GPI-anchor;
KW 3d-structure.
FT SIGNAL 1 19 NEURAL CELL ADHESION MOLECULE 1, 120 kDa
FT CHAIN 20 725 ISOFORM,
FT DOMAIN 20 111 IG-DIXE C2-TYPE 1.
FT DOMAIN 116 205 IG-LIKE C2-TYPE 2.
FT DOMAIN 212 302 IG-LIKE C2-TYPE 3.
FT DOMAIN 309 402 IG-LIKE C2-TYPE 4.
FT DOMAIN 407 492 IG-LIKE C2-TYPE 5.
FT DOMAIN 519 596 FIBRONECTIN TYPE-IIII 1.
FT DOMAIN 625 692 FIBRONECTIN TYPE-IIII 2.
FT DOMAIN 152 156 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 161 165 HEPARIN-BINDING (POTENTIAL).
FT DISUFID 41 96 PROBABLE.
FT DISUFID 139 189 PROBABLE.
FT DISUFID 235 288 PROBABLE.
FT DISUFID 330 386 PROBABLE.
FT DISUFID 427 480 PROBABLE.
FT CARBOHYD 222 222 N-LINKED (GLCNAC...)
FT CARBOHYD 316 316 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 348 348 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 424 424 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 479 479 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 261 268 ERSSSVTS->DEKHFSFD (IN REF. 2).
FT CONFLICT 273 273 V-> L (IN REF. 2).
FT CONFLICT 354 355 QD-> K (IN REF. 2).
FT CONFLICT 549 549 T-> R (IN REF. 2).
FT CONFLICT 572 572 D-> V (IN REF. 2).
FT CONFLICT 575 575 D-> R (IN REF. 2).
FT CONFLICT 589 594 MORSES-> SAATRF (IN REF. 2).
FT CONFLICT 600 602 PEL-> REP (IN REF. 2).
FT CONFLICT 657 657 H-> D (IN REF. 2 AND 3).
SQ SEQUENCE 725 AA; 80296 MW; G2AEBSB446IC6B2F CRC64;

| | | | | |
|-----------------------|------------------|-----------------|-------------|-------------|
| Query Match | 5.4%; | Score 125.5; | DB 1; | Length 725; |
| Best Local Similarity | 20.9%; | Pred. No. 0.41; | | |
| Matches 95; | Conservative 64; | Mismatches 187; | Indels 109; | Gaps 21; |

```

Oy      RG-VPRHLLLVQLALLPAATQGNKVVLGKKGGPVEVLTCTASQKKSIQFHM-KNSNQIK 60
      || : : : : :
Db      196 RGEINFEDIOVIYVNPVTAQKOSIYNATANIQGSVTLVCDADGPEPEPTMSWTKDGEPIE 255
      || : : : : :
Oy      61 ILIQGSPFLTKGPKSLNDRAD--SRRLSDQGNFPLIINKLIEDSDTYICEVE---DQ 114
      || : : : : :
Db      256 -----NEEDERSRSVSADSE--VTIRVNDKNDDEAYVCIANKAGEQ 297
      || : : : : :
Oy      115 KEVQVLVFG---LTANSDTHLQGSGLTLTLBSPSSPSVQCR-----SRGNIQ 164
      || : : : : :
Db      298 DASIHKLVPFAKPIITYENQTALEEOVTLTICEASGPISITWIRSTRNISSEODLD 357
      || : : : : :
Oy      165 GG-----KTLISOLELDQSGTWTCTVLAQNKVEFKIDIVLAFQKASSI----- 210
      || : : : : :
Db      358 GHVWVSHARVSSLTLSKSIQYRDAGSYMTASNTIGQSGSID---LEFQVAPLQSPVA 414
      || : : : : :
Oy      211 VYKKEGEQVE---FSPLAFTVEKLTGSGELMQAERASSSKSWITFDLKN-KEYSVK 264
      || : : : : :
Db      415 VYTWEGQVWITCEVAFPSA-TIS-----WFRDQLPLSSNYSNIKIYNTPSASYL 465
      || : : : : :
Oy      265 RYVQDPKLQMGK-----KLPLHLTLPOA-----LPQVAGSGNLTLLALEAK 304
      || : : : : :
Db      466 EYVTPDSNDPFGNYNCTAIVNRIGOSILEFTLVADTPSSPSIDRKEPPSSITQVOPDEPEA 525
      || : : : : :
Oy      305 TGLKHQEVNLVYVNRATQKNLTICEWGPSTSPKMLSLKLENKEAKVSKREKPPVYVNLPE 364
      || : : : : :
Db      526 TG-----GVPIIKYAEWKSIGEEBSMHT-----WYDKAEANMGIVL-----IMGLKPE 570
      || : : : : :
Oy      365 AGMQCCLSDSGQVL-----LESNTKVLPTWSTP 393
      || : : : : :
Db      571 TTVSDRLAALNGKGLGEIWPQSESKTQVPELSAP 605

```

| RESULT 32 | DCC_HUMAN | STANDARD; | PRT; 1447 AA. |
|-----------|--|-----------|---------------|
| ID | DCC_HUMAN | | |
| AC | P43146; | | |
| DT | 01-NOV-1995 (Rel. 32, Created) | | |
| DT | 01-NOV-1995 (Rel. 32, Last sequence update) | | |
| DT | 10-OCT-2003 (Rel. 42, Last annotation update) | | |
| DE | Tumor suppressor protein DCC precursor (colorectal cancer suppressor). | | |
| GN | DCC. | | |
| OS | Homo sapiens (Human). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| OX | NCBI Taxid=9606; | | |
| LN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RX | MEDLINE=95011532; PubMed=7926722; | | |
| RA | Hedrick L., Cho K.R., Fearon E.R., Wu T.-C., Kinzler K.W., | | |
| RA | "the DCC gene product in cellular differentiation and colorectal | | |
| RT | tumorigenesis."; | | |
| RT | Genes Dev. 8:1174-1183(1994). | | |
| LN | [2] | | |
| RP | SEQUENCE OF 1-750 FROM N.A. | | |
| RX | MEDLINE=90100559; PubMed=2294591; | | |
| RA | Fearon E.R., Cho K.R., Ngirto J.M., Kern S.E., Simons J.M., | | |
| RA | Ruppert J.M., Hamilton S.R., Frelisinger A.C., Thomas G., Kinzler K.W., | | |
| RA | Vogelstein B., | | |
| RT | "identification of a chromosome 18q gene that is altered in | | |
| RT | colorectal cancers."; | | |
| RL | Science 247:49-56(1990). | | |
| LN | [3] | | |
| RP | SEQUENCE OF 107-472 FROM N.A. (SCRAMBELD EXONS). | | |
| RX | MEDLINE=91121517; PubMed=1991332; | | |
| RA | Ngirto J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M., | | |
| RA | Oliver J.D., Kinzler K.W., Vogelstein B., | | |

```

FT DOMAIN 1123 1447 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 26 135 IG-LIKE C2-TYPE 1.
FT DOMAIN 139 229 IG-LIKE C2-TYPE 2.
FT DOMAIN 234 326 IG-LIKE C2-TYPE 3.
FT DOMAIN 331 416 IG-LIKE C2-TYPE 4.
FT DOMAIN 426 522 FIBRONECTIN TYPE-III 1.
FT DOMAIN 525 618 FIBRONECTIN TYPE-III 2.
FT DOMAIN 619 716 FIBRONECTIN TYPE-III 3.
FT DOMAIN 722 816 FIBRONECTIN TYPE-III 4.
FT DOMAIN 840 940 FIBRONECTIN TYPE-III 5.
FT DOMAIN 941 1042 FIBRONECTIN TYPE-III 6.
FT DISULFID 61 117 BY SIMILARITY.
FT DISULFID 161 212 BY SIMILARITY.
FT DISULFID 261 310 BY SIMILARITY.
FT DISULFID 352 400 BY SIMILARITY.
FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 702 702 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 168 168 M -> T (in oesophageal carcinoma).
FT VARIANT 201 201 R -> G.
FT VARIANT 201 201 P -> H (in a colorectal carcinoma).
FT VARIANT 1375 1375 /FTID=VAR_003911.
FT CONFLICT 138 138 MISSING (IN REF. 3).
FT CONFLICT 233 329 MISSING (IN REF. 3).
FT CONFLICT 421 421 MISSING (IN REF. 3).
SQ SEQUENCE 1447 AA; 158456 MW; 4A861276ED0471F CRC64;

```

```

Query Match 5.4%; Score 125.5; DB 1; Length 1447;
Beet Local Similarity 19.3%; Pred. No. 1;
Matches 77; Conservative 46; Mismatches 129; Indels 147; Gaps 15;

```

```

QY 34 GDTVELTCTASQKSIQPHWKNNOIKILGNQSGFETKPSKLTNRADRSRLMDQGNP 93
DB 154 GDTVLKCEVIGEPMTTHMKNQO-----DITP-----GDSRVVVLPSG--A 196
QY 94 LIKNLKIEDSDTYICEVED-----QKEVQL-----VEGLTANSDTHLQSQS 138
DB 197 LQISRLQPDIGIYCSANPNASRTGNAEVRILSDPOLHRLQVFLQRPSPVVAIEGKD 256
QY 139 LTL---TLSPPGS-----SPSVQCRSPRGKNIOGGKTLVSQLELDSCWTCTVION 189
DB 257 AVLCECVSGPPSPFTWLGRGEVILDRSKK-YSLGGSNLILSNVTDDSGMYTCVVYK 315
QY 190 OKKVEFKIDIVLA---FOKASSIYKKKEGEVEFSPFLAFVEXLTGSGELMWQAEAS 246
DB 316 NENISASAEITVLPWPFLNPSNITYAESMDIER-----ECTVSG-----KPV 359
QY 247 SSKSWITPDLNKKEYSVKRVTDPKLQWKKLPLHLTLFQALPOYAGSGNLTALAEAKTG 306
DB 360 PTVNM-----MKNGDV-----VIPSDFQIVGGSNLAI----- 387
QY 307 KLHQEVNLYVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPVVNLNPEAG 366
DB 388 -----LGIVKSDSG 396
QY 367 MWQCLSDSGQVLESINIKVLPFTWSPVHPRASALPAPP 405
DB 397 FYQCVANENAGNAGNOTSQQLIVP---KPAIPSSSVLPSAP 432

```

```

RESULT 33
PGM_HUMAN STANDARD; PRT; 4391 AA.
AC P98160; Q16287; G9H3V5;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Basement membrane-specific heparan sulfate proteoglycan core

```

```

DE protein precursor (HSPG) (Perlecan) (PLC).
GN HSPG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92112994; PubMed=1730768;
RA Kallunki P., Tryggvason K.;
RT "Human basement membrane heparan sulfate proteoglycan core protein: a
RT 467-kD protein containing multiple domains resembling elements of the
RT low density lipoprotein receptor, laminin, neural cell adhesion
RT molecules, and epidermal growth factor.";
RL J. Cell Biol. 116:559-571(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92235084; PubMed=1569102;
RA Murdoch A.D., Dodge G.R., Cohen I., Tuan R.S., Iozzo R.V.;
RT "Primary structure of the human heparan sulfate proteoglycan from
RT basement membrane (HSPG2/perlecan). A chimeric molecule with multiple
RT domains homologous to the low density lipoprotein receptor, laminin,
RT neural cell adhesion molecules, and epidermal growth factor.";
RL J. Biol. Chem. 267:8544-8557(1992).
RN [3]
RP SEQUENCE OF 22-4391 FROM N.A., AND VARIANT SJS1 TYR-1532.
RX MEDLINE=20553141; PubMed=11101850;
RA Nicole S., Davoine C.-S., Topaloglu H., Cattoiello L., Barracl D.,
RA Beighton P., Ben-Hamida C., Hammouda H., Cruaud C., White P.S.,
RA Sanson D., Urtizberea J.A., Lehmann-Horn F., Weissenbach J.,
RA Hentati F., Fontaine B.;
RT "Perlecan, the major proteoglycan of basement membranes, is altered in
RT patients with Schwartz-Jampel syndrome (Chondrodysdystrophic myotonia).";
RL Nat. Genet. 26:480-483(2000).
RN [4]
RP SEQUENCE OF 1016-1470 FROM N.A.
RX MEDLINE=91365376; PubMed=1679749;
RA Dodge G.R., Kovalszky I., Chu M.L., Hassell J.R., McBride O.W.,
RA Yi H.F., Iozzo R.V.;
RT "Heparan sulfate proteoglycan of human colon: partial molecular
RT cloning, cellular expression, and mapping of the gene (HSPG2) to the
RT short arm of human chromosome 1.";
RL Genomics 10:673-680(1991).
RN [5]
RP SEQUENCE OF 890-1396 FROM N.A.
RX MEDLINE=92120660; PubMed=1685141;
RA Kallunki P., Eddy R.L., Byers M.G., Kestila M., Shows T.B.,
RA Tryggvason K.;
RT "Cloning of human heparan sulfate proteoglycan core protein,
RT assignment of the gene (HSPG2) to 1p36.1-3p35 and identification of
RT a BamHI restriction fragment length polymorphism.";
RL Genomics 11:389-396(1991).
RN [6]
RP SEQUENCE OF 1-21 FROM N.A.
RX MEDLINE=94052171; PubMed=8234307;
RA Cohen I.R., Graessel S., Murdoch A.D., Iozzo R.V.;
RT "Structural characterization of the complete human perlecan gene and
RT its promoter.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:10404-10408(1993).
RN [7]
RP CARBOHYDRATE-LINKAGE SITE ASN-2121.
RX MEDLINE=22660472; PubMed=12754519;
RA Zhang H., Li X.-Y., Martin D.B., Aebbersold R.;
RT "Identification and quantification of N-linked glycoproteins using
RT hydrazide chemistry: stable isotope labeling and mass spectrometry.";
RL Nat. Biotechnol. 21:660-666(2003).
RN [8]
RP FUNCTION: This protein is an integral component of basement
RN membranes. It is responsible for the fixed negative electrostatic
RN charge and is involved in the charge-selective ultrafiltration
RN properties. It serves as an attachment substrate for cells.

```

```
CC -1- SUBUNIT: Purified perlecan has a strong tendency to aggregate in
CC dimers or stellate structures. It interacts with other basement
CC membrane components such as laminin, prolargin and collagen type
CC IV.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: Found in the basement membranes.
CC -1- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
CC AND O-LINKED OLIGOSACCHARIDES.
CC DISEASE: Defects in HSPG2 are the cause of Schwartz-Jampel
CC syndrome (SJS1) [MIM:255800]; a rare autosomal recessive disorder
CC characterized by permanent myotonia (prolonged failure of muscle
CC relaxation) and skeletal dysplasia, resulting in reduced stature,
CC kyphoscoliosis, bowing of the diaphyses and irregular epiphyses.
CC -1- SIMILARITY: Contains 4 LDL-receptor class A domains.
CC -1- SIMILARITY: Contains 11 laminin EGF-like domains.
CC -1- SIMILARITY: Contains 3 laminin IV domains.
CC -1- SIMILARITY: Contains 22 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 3 laminin G-like domains.
CC -1- SIMILARITY: Contains 4 EGF-like domains.
CC -1- SIMILARITY: Contains 1 SEA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See license@isb-sib.ch).
CC -----
DR EMBL; X62515; CAA44373.1; -.
DR EMBL; M85289; AAA52700.1; -.
DR EMBL; AL445795; CAC18534.1; -.
DR EMBL; M64283; AAA52699.1; -.
DR EMBL; S76436; AAB21121.2; -.
DR EMBL; L22078; -. NOT_ANNOTATED_CDS.
DR PIR; A38096; A38096.
DR HSSP; P00740; 1EDM.
DR Genea-2DPAGE; P96160; -.
DR Genea; HGNC:5273; HSPG2.
DR MIM; 142461; -.
DR MIM; 255800; -.
DR InterPro; IPR006985; CONA_1like_jec_gl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_1like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR003596; IG_V.
DR InterPro; IPR000034; Laminin_B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000082; SEA_domain.
DR Pfam; PF00008; EGF_4.
DR Pfam; PF00047; IG_22.
DR Pfam; PF00052; laminin_B_3.
DR Pfam; PF00053; laminin_EGF_7.
DR Pfam; PF00054; laminin_G_3.
DR Pfam; PF00057; ldl_recept_a; 4.
DR Pfam; PF01390; SEA; 1.
DR PRINTS; PR00261; LDLRECEPTOR.
DR PRODOM; PD003031; Laminin_B_3.
DR SMART; SM00181; EGF_15.
DR SMART; SM00180; EGF_Lam; 12.
DR SMART; SM00409; IG_22.
DR SMART; SM00408; IG_C2; 21.
DR SMART; SM00406; IGV_7.
DR SMART; SM00281; Lamb; 3.
DR SMART; SM00282; Lamb; 3.
DR SMART; SM00192; LDLA; 4.
DR SMART; SM00200; SEA; 1.
DR PROSITE; PS00022; EGF_1; 9.
```

| Query Match | Best Local Similarity | Matches | Score | DB 1 | Length | DB 2 | Score | DB 3 | Length | DB 4 | Score | DB 5 | Length | DB 6 | Score | DB 7 | Length | DB 8 | Score | DB 9 | Length | DB 10 | Score | DB 11 | Length | DB 12 | Score | DB 13 | Length | DB 14 | Score | DB 15 | Length | DB 16 | Score | DB 17 | Length | DB 18 | Score | DB 19 | Length | DB 20 | Score | DB 21 | Length | DB 22 | Score | DB 23 | Length | DB 24 | Score | DB 25 | Length | DB 26 | Score | DB 27 | Length | DB 28 | Score | DB 29 | Length | DB 30 | Score | DB 31 | Length | DB 32 | Score | DB 33 | Length | DB 34 | Score | DB 35 | Length | DB 36 | Score | DB 37 | Length | DB 38 | Score | DB 39 | Length | DB 40 | Score | DB 41 | Length | DB 42 | Score | DB 43 | Length | DB 44 | Score | DB 45 | Length | DB 46 | Score | DB 47 | Length | DB 48 | Score | DB 49 | Length | DB 50 | Score | DB 51 | Length | DB 52 | Score | DB 53 | Length | DB 54 | Score | DB 55 | Length | DB 56 | Score | DB 57 | Length | DB 58 | Score | DB 59 | Length | DB 60 | Score | DB 61 | Length | DB 62 | Score | DB 63 | Length | DB 64 | Score | DB 65 | Length | DB 66 | Score | DB 67 | Length | DB 68 | Score | DB 69 | Length | DB 70 | Score | DB 71 | Length | DB 72 | Score | DB 73 | Length | DB 74 | Score | DB 75 | Length | DB 76 | Score | DB 77 | Length | DB 78 | Score | DB 79 | Length | DB 80 | Score | DB 81 | Length | DB 82 | Score | DB 83 | Length | DB 84 | Score | DB 85 | Length | DB 86 | Score | DB 87 | Length | DB 88 | Score | DB 89 | Length | DB 90 | Score | DB 91 | Length | DB 92 | Score | DB 93 | Length | DB 94 | Score | DB 95 | Length | DB 96 | Score | DB 97 | Length | DB 98 | Score | DB 99 | Length | DB 100 | Score | DB 101 | Length | DB 102 | Score | DB 103 | Length | DB 104 | Score | DB 105 | Length | DB 106 | Score | DB 107 | Length | DB 108 | Score | DB 109 | Length | DB 110 | Score | DB 111 | Length | DB 112 | Score | DB 113 | Length | DB 114 | Score | DB 115 | Length | DB 116 | Score | DB 117 | Length | DB 118 | Score | DB 119 | Length | DB 120 | Score | DB 121 | Length | DB 122 | Score | DB 123 | Length | DB 124 | Score | DB 125 | Length | DB 126 | Score | DB 127 | Length | DB 128 | Score | DB 129 | Length | DB 130 | Score | DB 131 | Length | DB 132 | Score | DB 133 | Length | DB 134 | Score | DB 135 | Length | DB 136 | Score | DB 137 | Length | DB 138 | Score | DB 139 | Length | DB 140 | Score | DB 141 | Length | DB 142 | Score | DB 143 | Length | DB 144 | Score | DB 145 | Length | DB 146 | Score | DB 147 | Length | DB 148 | Score | DB 149 | Length | DB 150 | Score | DB 151 | Length | DB 152 | Score | DB 153 | Length | DB 154 | Score | DB 155 | Length | DB 156 | Score | DB 157 | Length | DB 158 | Score | DB 159 | Length | DB 160 | Score | DB 161 | Length | DB 162 | Score | DB 163 | Length | DB 164 | Score | DB 165 | Length | DB 166 | Score | DB 167 | Length | DB 168 | Score | DB 169 | Length | DB 170 | Score | DB 171 | Length | DB 172 | Score | DB 173 | Length | DB 174 | Score | DB 175 | Length | DB 176 | Score | DB 177 | Length | DB 178 | Score | DB 179 | Length | DB 180 | Score | DB 181 | Length | DB 182 | Score | DB 183 | Length | DB 184 | Score | DB 185 | Length | DB 186 | Score | DB 187 | Length | DB 188 | Score | DB 189 | Length | DB 190 | Score | DB 191 | Length | DB 192 | Score | DB 193 | Length | DB 194 | Score | DB 195 | Length | DB 196 | Score | DB 197 | Length | DB 198 | Score | DB 199 | Length | DB 200 | Score | DB 201 | Length | DB 202 | Score | DB 203 | Length | DB 204 | Score | DB 205 | Length | DB 206 | Score | DB 207 | Length | DB 208 | Score | DB 209 | Length | DB 210 | Score | DB 211 | Length | DB 212 | Score | DB 213 | Length | DB 214 | Score | DB 215 | Length | DB 216 | Score | DB 217 | Length | DB 218 | Score | DB 219 | Length | DB 220 | Score | DB 221 | Length | DB 222 | Score | DB 223 | Length | DB 224 | Score | DB 225 | Length | DB 226 | Score | DB 227 | Length | DB 228 | Score | DB 229 | Length | DB 230 | Score | DB 231 | Length | DB 232 | Score | DB 233 | Length | DB 234 | Score | DB 235 | Length | DB 236 | Score | DB 237 | Length | DB 238 | Score | DB 239 | Length | DB 240 | Score | DB 241 | Length | DB 242 | Score | DB 243 | Length | DB 244 | Score | DB 245 | Length | DB 246 | Score | DB 247 | Length | DB 248 | Score | DB 249 | Length | DB 250 | Score | DB 251 | Length | DB 252 | Score | DB 253 | Length | DB 254 | Score | DB 255 | Length | DB 256 | Score | DB 257 | Length | DB 258 | Score | DB 259 | Length | DB 260 | Score | DB |
|-------------|-----------------------|---------|-------|------|--------|------|-------|------|--------|------|-------|------|--------|------|-------|------|--------|------|-------|------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|----|
|-------------|-----------------------|---------|-------|------|--------|------|-------|------|--------|------|-------|------|--------|------|-------|------|--------|------|-------|------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|-------|-------|-------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|--------|--------|--------|-------|----|

```

OY 290 QY---AGSG---NLTLAEAKTGKTLHQEVLNVM---RATOL----- 322
DB 3391 ATSIAGSPTVQVPTOLETKSIGASVEFHCAVPSDRGQLWFKEGGLPRGHVQDGV 3450
OY 323 -----QKMLTCEVWGSP-----TSPKMLSLKLENKAKV----- 351
DB 3451 LRIONLDQSCQGTGYICQAHGPMWKAQASALVIALPVLINIRTSVQVNVGAHVEFEC 3510
OY 352 ----SKREKPVV-----LNP-----EAGMWQCLLSIS-----GQYLL 380
DB 3511 IALGDPKPVWTSKVGGHLRPGIIVOSGGVRIAHVELADAGQYRCTATNAAGTQSHVLL 3570
OY 381 ESNIKVLPWSTPVHPRASGALPAPPTGSALPDPGTASALPDP 422
DB 3571 --LVQALPQISMPQEVV-----PAGSAAVFPCLASGPTP 3604

RESULT 34
LACH SCHAM STANDARD; PRT; 349 AA.
AC 026474;
DB 15-JUL-1998 (Rel. 36, Created)
DB 15-JUL-1998 (Rel. 36, Last sequence update)
DB 10-OCT-2003 (Rel. 42, Last annotation update)
DE Lachesin precursor.
GN LAC.
OS Schistocerca americana (American grasshopper).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridoidea;
OC Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.
OX NCBI_TaxID=7009;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RA MEDLINE=94038693; PubMed=8223276;
RA Karistrom R.O., Wilder L.P., Bastiani M.J.;
RT "Lachesin: an immunoglobulin superfamily protein whose expression
RT correlates with neurogenesis in grasshopper embryos.";
RL Development 118:509-522(1993).
CC -1- FUNCTION: MAY PLAY A ROLE IN EARLY NEURONAL DIFFERENTIATION AND
CC AXON OUTGROWTH.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC ONLY THOSE CELLS THAT BECOME NEUROBLASTS CONTINUE TO EXPRESS IT.
CC EXPRESSED BY NEUROBLASTS, GANGLION MOTHER CELLS AND NEURONS EARLY
CC IN THEIR LIVES, BUT EXPRESSION BECOMES RESTRICTED TO A SUBSET OF
CC NEURONS AS DEVELOPMENT PROGRESSES. EXPRESSED BY SENSORY NEURONS AS
CC THEY DELAMINATE FROM THE BODY WALL ECTODERM. IT IS ALSO PRESENT ON
CC GROWING AXONS OF THE CNS AND PNS AND BECOMES RESTRICTED TO A
CC SUBSET OF AXONS LATER IN DEVELOPMENT.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED ON DIFFERENTIATING NEURONAL CELLS
CC FROM THE ONSET OF NEUROGENESIS IN BOTH THE CENTRAL AND PERIPHERAL
CC NERVOUS SYSTEMS.
CC -1- PTM: The N-terminus is blocked.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; LJ3256; AAC37185.1; -.
CC DR HSSP; P80362; IMTL.
CC DR InterPro; IPR007110; Ig-like.
CC DR InterPro; IPR003598; Ig_C2.
CC DR Pfam; PF00047; Ig_3.
CC DR SMART; SM00408; IgC2; 2.
CC DR PROSITE; PS50835; IG_LIKE; 3.
CC DR Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
CC Repeat; Signal; Lipoprotein.

```

```

FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 332 LACHESIN.
FT PROPEP 333 349 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 22 127 IG-LIKE V-TYPE.
FT DOMAIN 132 218 IG-LIKE C2-TYPE 1.
FT DOMAIN 222 315 IG-LIKE C2-TYPE 2.
FT DISULFID 43 110 POTENTIAL.
FT DISULFID 154 201 POTENTIAL.
FT DISULFID 244 299 POTENTIAL.
FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 332 332 GPI-anchor amidated glycine (Potential).
SQ SEQUENCE 349 AA; 38974 MM; 5F13944BF849689 CRC64;

Query Match 5.4%; Score 125; DB 1; Length 349;
Beet Local Similarity 23.3%; Pred. No. 0.17; Indels 28; Gaps 7;
Matches 62; Conservative 39; Mismatches 137;

OY 34 GDTVELCTASQKSIQPHW----KNSNQIKILNQGSGFLTKGSPKLNDRASRRSLWDQ 89
DB 36 GGTVELECSVOYADYPTLMKMKVDRNQVDLPISGSSLIIRSRFALRYDTASS----- 91
OY 90 GNFLIINLKIEDSDTYICEV----EDQKEVQLVFG--LTNSDTHLLOGQSITL 141
DB 92 -TYLQIKDIDETDAGFYQCQCVIIGLNKKITAEVLDQVRPVPVISDSTRSLVSEGOAV 150
OY 142 TLESPPGS--SPSVQGRS-----PRGNIGGKTLISQSOELDSDGHWCTCTVLQNQKV 193
DB 151 RLECYAGGVPAPRVSRWRENNAILPTGSIYRGVNLKSRIGKEDRGTYVCVAENGVGK 210
OY 194 EFKIDIVLAFQKSSIVYKKEGQVERSPFLAETVETLTSGLMWAQERASSSKSWIT 253
DB 211 ARRNIAVEVERPPVITVPRRPLGQALQYDMLCEHVEVAPPATWLKDEIVLSNQHYS 270
OY 254 --FDLKNKEVSVKRVTDPLQMGK 276
DB 271 ISHPATDEFTDTRVITIERKQYQK 296

RESULT 35
DSCA HUMAN STANDARD; PRT; 2012 AA.
AC 060469; OS0468;
DB 16-OCT-2001 (Rel. 40, Created)
DB 16-OCT-2001 (Rel. 40, Last sequence update)
DB 10-OCT-2003 (Rel. 42, Last annotation update)
DE Down syndrome cell adhesion molecule precursor (CHD2).
GN DSCAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Brain;
RX MEDLINE=98087574; PubMed=9426258;
RA Yamakawa K., Huot Y.-K., Haendelt M.A., Hubert R., Chen X.-N.,
RA Lyons G.E., Korenberg J.R.;
RT "DSCAM: a novel member of the immunoglobulin superfamily maps in a
RT Down syndrome region and is involved in the development of the
RT nervous system.";
RL Hum. Mol. Genet. 7:227-237(1998).
RN [2]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=20384934; PubMed=10925149;
RA Agatawala K.L., Nakamura S., Tsutsumi Y., Yamakawa K.;
RT "Down syndrome cell adhesion molecule DSCAM mediates homophilic
RT intercellular adhesion.";
RL Brain Res. Mol. Brain Res. 79:118-126(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Tocioki Y., Choi D.-K., Groner Y.,

```

| | |
|----|---|
| RA | Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blehschmidt K., Pooley A., Menzel U., Delater T., Sasaki K., Lehmann R., Patterson D., Reichwald K., Kump A., Schillnabel M., Shudy A., Zimmermann W., Roentzsch A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E., Minoshima S., Shimizu N., Nordtief G., Hornischer K., Brandt P., Rascher F., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H., Rameyer J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E., Rameyer S., Borzym K., Gardiner K., Nizetic D., Francis F., Lehmach H., Reinhardt R., Yaspo M.-L.; |
| RT | "The DNA sequence of human chromosome 21."; |
| RL | Nature 405:311-319 (2000). |
| CC | - FUNCTION: CELL ADHESION MOLECULE THAT CAN MEDIATE CATION-INDEPENDENT HOMOPHILIC BINDING ACTIVITY. COULD BE INVOLVED IN NEUROUS SYSTEM DEVELOPMENT. |
| CC | - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PROBABLE). THE SHORT ISOFORM MAY BE SECRETED. |
| CC | - ALTERNATIVE PRODUCTS: |
| CC | Event=Alternative splicing; Named isoforms=2; |
| CC | Name=Long; Synonyms=CHD2-52; |
| CC | Isoid=O60469-1; Sequence=Displayed; |
| CC | Name=Short; Synonyms=CHD2-42; |
| CC | Isoid=O60469-2; Sequence=VSP_002502, VSP_002503; |
| CC | - TISSUE SPECIFICITY: Primarily expressed in brain. |
| CC | - SIMILARITY: Contains 6 immunoglobulin-like C2-type domains. |
| CC | - SIMILARITY: Contains 6 fibronectin type III domains. |
| CC | ----- |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). |
| CC | ----- |
| DR | EMBL; AF0234450; AAC17967.1; - |
| DR | EMBL; AF0234449; AAC17966.1; - |
| DR | EMBL; AF217525; AAF27525.1; - |
| DR | EMBL; AL163283; CAB90464.1; - |
| DR | EMBL; AL163282; CAB90436.1; - |
| DR | EMBL; AL163281; CAB90444.1; - |
| DR | Genew; HGNC:3039; DSCAM. |
| MM | 602523; |
| DR | GO; GO:0005887; C:integral to plasma membrane; TAS. |
| DR | GO; GO:0005624; C:membrane fraction; TAS. |
| DR | GO; GO:0007155; P:cell adhesion; TAS. |
| DR | GO; GO:0007399; P:neurogenesis; TAS. |
| DR | InterPro; IPR008957; FN III-like. |
| DR | InterPro; IPR003961; FN III. |
| DR | InterPro; IPR007110; Ig-like. |
| DR | InterPro; IPR003598; Ig_c2. |
| DR | Pfam; PF00041; fn3; 6. |
| DR | Pfam; PF00047; Ig; 9. |
| DR | SMART; SM00060; FN3; 6. |
| DR | SMART; SM00408; IGC2; 7. |
| DR | PROSITE; PS00835; IG-LIKE; 9. |
| KW | immunoglobulin domain; Glycoprotein; Signal; Cell adhesion; Repeat; Transmembrane; Alternative splicing. |
| FT | SIGNAL 1 |
| FT | CHAIN 17 |
| FT | DOWN 18 |
| FT | DOMAIN 18 |
| FT | TRANSMEM 1596 |
| FT | POTENTIAL 1616 |
| FT | CYTOPLASMIC (POTENTIAL) 1617 |
| FT | DOMAIN 129 |
| FT | IG-LIKE C2-TYPE 1 |
| FT | IG-LIKE C2-TYPE 2 |
| FT | IG-LIKE C2-TYPE 3 |
| FT | IG-LIKE C2-TYPE 4 |
| FT | IG-LIKE C2-TYPE 5 |
| FT | IG-LIKE C2-TYPE 6 |
| FT | IG-LIKE C2-TYPE 7 |
| FT | IG-LIKE C2-TYPE 8 |
| FT | IG-LIKE C2-TYPE 9 |
| FT | IG-LIKE C2-TYPE 10 |
| FT | IG-LIKE C2-TYPE 11 |
| FT | IG-LIKE C2-TYPE 12 |
| FT | IG-LIKE C2-TYPE 13 |
| FT | IG-LIKE C2-TYPE 14 |
| FT | IG-LIKE C2-TYPE 15 |
| FT | IG-LIKE C2-TYPE 16 |
| FT | IG-LIKE C2-TYPE 17 |
| FT | IG-LIKE C2-TYPE 18 |
| FT | IG-LIKE C2-TYPE 19 |
| FT | IG-LIKE C2-TYPE 20 |
| FT | IG-LIKE C2-TYPE 21 |
| FT | IG-LIKE C2-TYPE 22 |
| FT | IG-LIKE C2-TYPE 23 |
| FT | IG-LIKE C2-TYPE 24 |
| FT | IG-LIKE C2-TYPE 25 |
| FT | IG-LIKE C2-TYPE 26 |
| FT | IG-LIKE C2-TYPE 27 |
| FT | IG-LIKE C2-TYPE 28 |
| FT | IG-LIKE C2-TYPE 29 |
| FT | IG-LIKE C2-TYPE 30 |
| FT | IG-LIKE C2-TYPE 31 |
| FT | IG-LIKE C2-TYPE 32 |
| FT | IG-LIKE C2-TYPE 33 |
| FT | IG-LIKE C2-TYPE 34 |
| FT | IG-LIKE C2-TYPE 35 |
| FT | IG-LIKE C2-TYPE 36 |
| FT | IG-LIKE C2-TYPE 37 |
| FT | IG-LIKE C2-TYPE 38 |
| FT | IG-LIKE C2-TYPE 39 |
| FT | IG-LIKE C2-TYPE 40 |
| FT | IG-LIKE C2-TYPE 41 |
| FT | IG-LIKE C2-TYPE 42 |
| FT | IG-LIKE C2-TYPE 43 |
| FT | IG-LIKE C2-TYPE 44 |
| FT | IG-LIKE C2-TYPE 45 |
| FT | IG-LIKE C2-TYPE 46 |
| FT | IG-LIKE C2-TYPE 47 |
| FT | IG-LIKE C2-TYPE 48 |
| FT | IG-LIKE C2-TYPE 49 |
| FT | IG-LIKE C2-TYPE 50 |
| FT | IG-LIKE C2-TYPE 51 |
| FT | IG-LIKE C2-TYPE 52 |
| FT | IG-LIKE C2-TYPE 53 |
| FT | IG-LIKE C2-TYPE 54 |
| FT | IG-LIKE C2-TYPE 55 |
| FT | IG-LIKE C2-TYPE 56 |
| FT | IG-LIKE C2-TYPE 57 |
| FT | IG-LIKE C2-TYPE 58 |
| FT | IG-LIKE C2-TYPE 59 |
| FT | IG-LIKE C2-TYPE 60 |
| FT | IG-LIKE C2-TYPE 61 |
| FT | IG-LIKE C2-TYPE 62 |
| FT | IG-LIKE C2-TYPE 63 |
| FT | IG-LIKE C2-TYPE 64 |
| FT | IG-LIKE C2-TYPE 65 |
| FT | IG-LIKE C2-TYPE 66 |
| FT | IG-LIKE C2-TYPE 67 |
| FT | IG-LIKE C2-TYPE 68 |
| FT | IG-LIKE C2-TYPE 69 |
| FT | IG-LIKE C2-TYPE 70 |
| FT | IG-LIKE C2-TYPE 71 |
| FT | IG-LIKE C2-TYPE 72 |
| FT | IG-LIKE C2-TYPE 73 |
| FT | IG-LIKE C2-TYPE 74 |
| FT | IG-LIKE C2-TYPE 75 |
| FT | IG-LIKE C2-TYPE 76 |
| FT | IG-LIKE C2-TYPE 77 |
| FT | IG-LIKE C2-TYPE 78 |
| FT | IG-LIKE C2-TYPE 79 |
| FT | IG-LIKE C2-TYPE 80 |
| FT | IG-LIKE C2-TYPE 81 |
| FT | IG-LIKE C2-TYPE 82 |
| FT | IG-LIKE C2-TYPE 83 |
| FT | IG-LIKE C2-TYPE 84 |
| FT | IG-LIKE C2-TYPE 85 |
| FT | IG-LIKE C2-TYPE 86 |
| FT | IG-LIKE C2-TYPE 87 |
| FT | IG-LIKE C2-TYPE 88 |
| FT | IG-LIKE C2-TYPE 89 |
| FT | IG-LIKE C2-TYPE |

[illegible]

```

OY 317 MRATOLQKLTCEVWGPTSPKMLSLKLEKKAQVSKREKPY--W-----VLNPEAGMW 368
DB 804 -----KEMSTARG-----EKPIYWKEDRIINPEMAY 835
OY 369 OCLLSDSQVLLSNIKVLPT 389
DB 836 LVSTKEVGEVY-STLQIILPT 855

RESULT 36
NCAL_MOUSE
ID NCAL_MOUSE STANDARD; PRT; 1115 AA.
AC P13595; Q61949;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neutral cell adhesion molecule 1, 180 kDa isoform precursor (N-CAM 180)
DE (NCAM-180).
GN NCAM1 OR NCAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM N-CAM 180).
RC STRAIN=C57BL/6;
RX MEDLINE=87246524; PubMed=3595653;
RA Barthele D., Santoni M.J., Wille W., Ruppert C., Chaix J.C.,
RA Hirsch M.R., Fontecilla-Camps J.C., Goridis C.;
RT "Isolation and nucleotide sequence of mouse NCAM cDNA that codes for
RT a Mr 79,000 polypeptide without a membrane-spanning region.";
RL EMBO J. 6:907-914(1987).
RN [2]
RP SEQUENCE OF 529-1115 FROM N.A. (ISOFORM N-CAM 140).
RC STRAIN=C57BL/6;
RX MEDLINE=88067687; PubMed=3684567;
RA Santoni M.-J., Barthele D., Barbae J.A., Hirsch M.-R., Steinmetz M.,
RA Goridis C., Wille W.;
RT "Analysis of cDNA clones that code for the transmembrane forms of the
RT mouse neural cell adhesion molecule (NCAM) and are generated by
RT alternative RNA splicing.";
RL Nucleic Acids Res. 15:8621-8641(1987).
RN [3]
RP SEQUENCE OF 642-1115 FROM N.A. (ISOFORM N-CAM 180).
RX MEDLINE=88283628; PubMed=3396534;
RA Barbae J.A., Chaix J.C., Steinmetz M., Goridis C.;
RT "Differential splicing and alternative polyadenylation generates
RT distinct NCAM transcripts and proteins in the mouse.";
RL EMBO J. 7:1625-632(1988).
RN [4]
RP SEQUENCE OF 804-1081 FROM N.A. (ISOFORM N-CAM 180).
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=88247737; PubMed=2454455;
RA Barthele D., Vopper G., Wille W.;
RT "NCAM-180, the large isoform of the neural cell adhesion molecule of
RT the mouse, is encoded by an alternatively spliced transcript.";
RL Nucleic Acids Res. 16:4217-4225(1988).
RN [5]
RP SEQUENCE OF 702-1115 FROM N.A. (ISOFORM N-CAM 140).
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=89251563; PubMed=2721486;
RA Santoni M.J., Barthele D., Vopper G., Boned A., Goridis C., Wille M.;
RT "Differential exon usage involving an unusual splicing mechanism
RT generates at least eight types of NCAM cDNA in mouse brain.";
RL EMBO J. 8:385-392(1989).
RN [6]
RP SEQUENCE OF 20-36.
RX MEDLINE=86140120; PubMed=3512556;
RA Rougon G., Marshak D.R.;
RT "Structural and immunological characterization of the amino-terminal
RT domain of mammalian neural cell adhesion molecules.";
RL J. Biol. Chem. 261:3396-3401(1986).
CC -1- FUNCTION: This protein is a cell adhesion molecule involved in

```

```

CC neuron-neuron adhesion, neurite fasciculation, outgrowth of
CC neurites, etc.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=N-CAM 180;
CC IsoId=P13595-1; Sequence=Displayed;
CC Name=N-CAM 140;
CC IsoId=P13595-2; Sequence=VSP_002588;
CC Name=N-CAM 120;
CC IsoId=P13594-1; Sequence=External;
CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X07200; CAA30177.1; -.
CC EMBL; X00051; -. NOT ANNOTATED_CDS.
CC EMBL; X06328; CAA29641.1; -.
CC EMBL; X07195; CAA30173.1; -.
CC EMBL; X07244; CAA30230.1; -.
CC EMBL; X15051; CAA33150.1; -.
CC EMBL; X15052; CAA33151.1; -.
CC PIR; A29673; IJMSNL.
CC MGI; MGI:97281; Ncam1.
CC InterPro; IPR008957; FN_III-like.
CC InterPro; IPR003961; FN_III.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC Pfam; PF00041; fn3; 2.
CC Pfam; PF00047; ig; 5.
CC SMART; SM00408; IGC2; 2.
CC SMART; SM00408; IGC2; 5.
CC PROSITE; PS50835; IG-LIKE; 5.
CC Cell adhesion; Glycoprotein; Transmembrane; Repeat;
CC Immunoglobulin domain; Alternative splicing; Signal; Heparin-binding.
CC FT SIGNAL 1 19
CC FT CHAIN 20 1115
CC FT 20 1115 NEURAL CELL ADHESION MOLECULE 1, 180 kDa
CC FT 20 1115 ISOFORM.
CC FT 20 1115 EXTRACELLULAR (POTENTIAL).
CC FT 712 729 POTENTIAL.
CC FT 729 729 CYTOPLASMIC (POTENTIAL).
CC FT 730 1115 IG-LIKE C2-TYPE 1.
CC FT 111 111 IG-LIKE C2-TYPE 1.
CC FT 116 205 IG-LIKE C2-TYPE 2.
CC FT 212 302 IG-LIKE C2-TYPE 3.
CC FT 302 402 IG-LIKE C2-TYPE 4.
CC FT 402 492 IG-LIKE C2-TYPE 5.
CC FT 492 519 FIBRONECTIN TYPE-III 1.
CC FT 519 596 FIBRONECTIN TYPE-III 2.
CC FT 596 625 FIBRONECTIN TYPE-III 2.
CC FT 625 692 HEPARIN-BINDING (POTENTIAL).
CC FT 692 152 HEPARIN-BINDING (POTENTIAL).
CC FT 152 156 HEPARIN-BINDING (POTENTIAL).
CC FT 156 165 HEPARIN-BINDING (POTENTIAL).
CC FT 165 165 HEPARIN-BINDING (POTENTIAL).
CC FT 41 96 PROBABLE.
CC FT 96 139 PROBABLE.
CC FT 139 189 PROBABLE.
CC FT 189 235 PROBABLE.
CC FT 235 288 PROBABLE.
CC FT 288 330 PROBABLE.
CC FT 330 386 PROBABLE.
CC FT 386 427 PROBABLE.
CC FT 427 480 PROBABLE.
CC FT 480 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT 522 566 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT 566 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT 616 666 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT 666 716 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT 716 766 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT 766 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT 816 879 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT 879 929 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT 929 979 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT 979 1029 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT 1029 1076 Missing (in isoform N-CAM 140).
CC FT 1076 /Ftd=VSP_002588.
CC SQ SEQUENCE 1115 AA; 119351 MW; 2C93DCD474CFBACF CRC64;
CC Query Match 5.3%; Score 124; DB 1; Length 1115;

```

```

Best Local Similarity 21.9%: Pred. No. 0.9; Mismatches 128; Indels 68; Gaps 15;
Matches 68; Conservative 47;

OY      3 RG-VEFRHLILVLQALLPAATQGNKVYLGGKKDVTVELTCTASQKKSIOFHV-KNSNQIK 60
        ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB     196 RGEINFKDIOIVLVNVPPTVGARQSIVNATANIAGSVTLVCADGPFPEPTMTXDKGEIE 255
OY      61 ILGNQGSFLTKGPSKLNDRAD--SRSLMDQGNFPLITKNLKIENSDPIYCEV---DQ 114
        -----NEEDERSRSSVS DSE--VTIRNVDKNDDEAYEVCIAENKAGEQ 297
DB     256 -----
OY      115 KEEVOLVFEGL--LTANSDFHLLQGSGSTLTLESPPSSPSYQCR-----SPRKNIQ 164
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB     298 DASILHKVFAPFKITYVENQTAMELEEOVTLTCESGDPISITWRTSTRNISSEQDL 357
OY      165 GG-----KTLSVSQLLELDQSGTWTCTVLQNOKKVEFKIDIVLAFOKASSI----- 210
        ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB     358 GHMVVRSHAVRSSLTKSIQRDAEWMTCTASNIGQSGSID---LEFGVAPLKQGPVA 414
OY      211 VYKKEGEQVE----FSFPPLAFYEKLTGSGELMWQAERASSSKSWTFIDLKN-KEYSVK 264
        :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB     415 VYTEMGNQVNLTICEVFAYPSA-TIS-----WPRDQLLPSSSNYSNIKTYNPASLYL 465
OY      265 RVTDOPRLQMG 275
        |||||:::|
DB     466 EVTPDSENDGF 476

RESULT 37
YTFN_HAEIN
ID_YTFN_HAEIN STANDARD: PRT: 1298 AA.
AC Q57523;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein HI0696.
GN HI0696.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
CX NCBI_TaxID=727;
RN RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischman R.D., Adams M.D., White
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Goeyne J.D.,
RA Scott J.D., Shtrey R., Liu L.-I., Glodex A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Ueberbach T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Frazer C.M., Smith H.O.,
RA Venner J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd."
RL Science 269:496-512(1995).
RN RN [2]
RP IDENTIFICATION BY MASS SPECTROMETRY.
RX MEDLINE=20137488; PubMed=10675023;
RA Langen H., Tacacs B., Evers S., Berndt P., Lahm H.W., WiPf B.,
RA Gray C., Fountoulakis M.;
RT "Two-dimensional map of the proteome of Haemophilus influenzae.";
RL Electrophoresis 21:411-429(2000).
CC -I- SIMILARITY: STRONG, TO E.COLI YTFN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
```

| Query Match | 5.3% | Score 124; | DB 1; | Length 1298; |
|-----------------------|--|-----------------|-------------|--------------|
| Best Local Similarity | 22.3% | Pred. No. 1.1; | | |
| Matches 109; | Conservative 58; | Mismatches 178; | Indels 144; | Gaps 22; |
| DC | EMBL; U32752; AAC22356.1; -. | | | |
| DR | PIR; A64157; A64157. | | | |
| DR | TIGR; HI0696; -. | | | |
| DR | InterPro; IPR007452; DUF490. | | | |
| DR | Pfam; PF04357; DUF490; 1. | | | |
| KM | Complete proteome. | | | |
| SC | SEQUENCE 1298 AA, 141222 MW, 2D905DBED9329B9 CRC64; | | | |
| QY | 7 FRHLVLVLQALPRATQGNKVLGGKQDVELTCTASQKKSIOFPHWKNSNQIKLGNQ | 66 | | |
| DB | 815 YRFEKLYIPKLTLLNADIQNNLVL-----KTDIVNH-----NQG | 848 | | |
| QY | 67 SFLRKSGKGNDRADSRSLMDQGNFLLIKNLKIDSPTYICEVED-QKEEYOLLVFG | 125 | | |
| DB | 849 RIV--GDHLDLDAKNR-----QGGTLAIERLNLSTANQLLTSGESVNGEVSFKLSFG | 900 | | |
| QY | 126 TANSDFHLLQG-----QSLLTLESPPG--SSPSVQCRSPKGNIOGGKTLSSV-SQLEI | 178 | | |
| DB | 901 -GNLEKPLNDPRIRIRITKLKMPNVIITDGDALAFNDNRSLQGGIKITVDHNLITG | 959 | | |
| QY | 179 SGTYT---CTVLNQKKEFKIDIVLAFOKAS-SIVYKKEGQVEFSPPLAFTEKLT | 233 | | |
| DB | 960 RANMANIEHMTTELNAQNNENVDIPMAKLRFPSPNITIKANPKEL-----NLS | 1008 | | |
| QY | 234 GSGGLMW-----QAERASSKSWITPDLKNKEVSVR-----VTQDPKL | 272 | | |
| DB | 1009 GTVVIDIPAKRIKIDSLPTAEVSSDEVILNGPHSKKEELIKREPAATKSGMEIRSDLRI | 1068 | | |
| QY | 273 QMGKKLPLHLTLPOALPOYAGSGULTALEAKTSK----LHOENVLVMRATQ-----L | 322 | | |
| DB | 1069 NIGQDV-----SLDAYGLKTLVDGLSLVKQDKNGLGFGQINLTIGKGRASFGDILLI | 1120 | | |
| QY | 323 QKNLTCEWGPSTSPKMLSLKLENKEAKVSKREKPVWLNE-----AGMNOCLLSD | 374 | | |
| DB | 1121 RKGILSFGAQIQPTL-----NIEA-----IRNDETMEDSKITGAVVIGIAD | 1163 | | |
| QY | 375 SGQVLESNIVKLPFTWTPVHPRASALPAPPTGSLALDPQTASALPDPAPASALPALAV | 434 | | |
| DB | 1164 SPEVTI-----FSEPSKPDQALSYLTIGRSLSSGEGVSTGSVGAA----- | 1205 | | |
| QY | 435 ISFLGLGL 443 | | | |
| DB | 1206 ---LIGLGI 1211 | | | |
| RESULT 38 | | | | |
| NCM2_HUMAN | | | | |
| ID_NCM2_HUMAN | STANDARD; | PRT; | 837 AA. | |
| AC | O15394; | | | |
| DT | 15-JUL-1998 (Rel. 36, Created) | | | |
| DT | 15-JUL-1998 (Rel. 36, Last sequence update) | | | |
| DT | 10-OCT-2003 (Rel. 42, Last annotation update) | | | |
| DE | Neural cell adhesion molecule 2 precursor (N-CAM 2). | | | |
| GN | NCAM2 OR NCAM21. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | | |
| OC | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Brain; | | | |
| RX | MEDLINE=97369930; PubMed=9226371; | | | |
| RA | Paciotti-Giacchino A., Chen H., Antonarakis S.E.; | | | |
| RT | "Cloning of a novel human neural cell adhesion molecule gene (NCAM2) | | | |
| RT | that maps to chromosome region 21q21 and is potentially involved in | | | |
| RT | Down syndrome."; | | | |
| RL | Genomics 43:43-51(1997). | | | |
| RN | [2] | | | |

RP CARBOHYDRATE-LINKAGE SITES ASN-445 AND ASN-562.
 RX MEDLINE=22660472; PubMed=12754519;
 RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
 RT "Identification and quantification of N-linked glycoproteins using
 hydrazide chemistry, stable isotope labeling and mass spectrometry,";
 RL Nat. Biotechnol. 21:660-666(2003).
 CC -1- FUNCTION: May play important roles in selective fasciculation and
 CC -1- zone-of-zone projection of the primary olfactory axons.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed most strongly in adult and fetal
 CC brain.
 CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U75330; AAB80803.1; -;
 CC Genew; HGNC:7657; NCAM2.
 DR MIM; 602040; -;
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR GO; GO:0005886; C:plasma membrane; TAS.
 DR GO; GO:0007158; P:neuronal cell adhesion; TAS.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003588; Ig_C2.
 DR Pfam; PF00041; FN3; 2.
 DR Pfam; PF00047; Ig; 5.
 DR SMART; SM00060; FN3; 2.
 DR SMART; SM00408; IGC2; 5.
 DR PROSITE; PS50835; IG_LIKE; 5.
 DR Cell adhesion, Transmembrane; Glycoprotein; Repeat;
 KW Immunoglobulin domain; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 837
 FT DOMAIN 20 697
 FT TRANSMEM 698 718
 FT DOMAIN 719 837
 FT DOMAIN 21 108
 FT DOMAIN 113 202
 FT DOMAIN 208 297
 FT DOMAIN 302 396
 FT DOMAIN 401 491
 FT DOMAIN 482 581
 FT DOMAIN 594 678
 FT DISULFID 42 93
 FT DISULFID 136 186
 FT DISULFID 232 281
 FT DISULFID 332 380
 FT DISULFID 422 475
 FT CARBOHYD 177 177
 FT CARBOHYD 219 219
 FT CARBOHYD 309 309
 FT CARBOHYD 406 406
 FT CARBOHYD 419 419
 FT CARBOHYD 445 445
 FT CARBOHYD 474 474
 FT CARBOHYD 562 562
 SQ SEQUENCE 837 AA; 92932 MW; C3D034106C5741C1 CRC64;

 Qy Best Match 5.3%; Score 123; DB 1; Length 837;
 Db Query Local Similarity 19.1%; Pred. No. 0.73; Indels 120; Gaps 24;
 Matches 92; Conservative 77; Mismatches 192;
 3 RG-VPEFHLLVLTALALPRAATGKGNVVGKKGDVETLCTASOKKSIOFHKNKSNQIKI 61
 193 REEIDRDIIVVNPVPAISMQKSPNATARGGEENTFSCRASGSPREPAISMFRNG--KL 250

Qy 62 LGNGSFLTKGPKSKLNDRAISRSLMDQGNFLIIKNLKIEDSPYICEV-----EDQKE 116
 Db 251 IENNEKYLKG-----SNTLTVRNIIINDGPFYVCRANKAGEDKQ 293
 Qy 117 E-VOLLVFGLTANSPDHLQ-----GSLTLTLESPPGSSPSVOC----- 155
 Db 294 AFLQYFV-----QPHLIQKNETTYENGQ-VTLVCBAEAEPIPEIWKKAVDGFTTE 345
 Qy 156 -RSPRG---KNIOGKTLVSQLELDOSGTWCTVLQ---NOKKVEFKIDIVLAFQ 205
 Db 346 GDKSPDGRIEVKGHGSSSLHAKDKVKSGRYCEASRIGCHQKMYLIEVPV-XFI 404
 Qy 206 KASSIVYKKEGEQVFEFSPPLAFVTEKLTGSELMWQARASSSKSWIFDLKNEKVSYKR 265
 Db 405 SNQITVYSWENPNINISCDV-----KSNPPASIHRRDKL-----VLPKN 445
 Qy 266 VTODPKLQMGKPLPHLTLPOLPYAGSGNLTALBAKTKLHQEVNLVVMRATOLQKN 325
 Db 446 TTNLKTSTGKMLEIA-PTSNDP-GRVYCT-----ATNHI---GTRPQEX 488
 Qy 326 LTCVWGFTSPKLMISLKLKLNKKAISKREK-----VVLINPEAGMQLIS 373
 Db 489 ILALADVPSSPYGVKIIELSGTTAKVS-FNRPDSHGVPVHHYQVDVKEVASIKWIVRS 547
 Qy 374 DSGVLESNIKVLPWSTPVPHPASALPAPPTG--SALPDPPTASAL-PDPASALPA 430
 Db 548 HGVQTVVVLN-NLEP--NTTYEIRVAANGKGGQYSKIEIFQTLPREPSPSHIGQS 604
 Qy 431 A 431
 Db 605 S 605

 RESULT 39
 ID C166 BRARE STANDARD; PRT; 564 AA.
 AC 090460; 090480;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE C166 antigen homolog precursor (Neuroilin) (DM-GRASP homolog).
 GN CD166.
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_Taxid=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94376084; PubMed=8089660;
 RA Kanki J.P., Chang S., Kuwada J.Y.;
 RT "The molecular cloning and characterization of potential chick
 RT DM-GRASP homologs in zebrafish and mouse.";
 RL J. Neurobiol. 25:831-845(1994).
 RN [2]
 RP SEQUENCE OF 398-561 FROM N.A.
 RA MEDLINE=94299040; PubMed=8026643;
 RA Laessle U., Giordano S., Stecher B., Lottspeich F., Stuermer C.A.O.;
 RT "Molecular characterization of fish neuroilin: a growth-associated
 RT cell surface protein and member of the immunoglobulin superfamily in
 RT the fish retinorectal system with similarities to chick protein
 RT DM-GRASP/SC-1/BEN.";
 RL Differentiation 56:21-29(1994).
 CC -1- FUNCTION: CELL ADHESION MOLECULE. INVOLVED IN NEURITE EXTENSION BY
 CC NEURONS VIA HETEROPHILIC AND HOMOPHILIC INTERACTIONS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like V-type domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

Qy 386 VLPTSTPVHPRASALPA--PPTGSAL 410
| | | | |
| | | | |

DR MIM; 110930; -.
DR GO; GO:0016021; C:integral to membrane; TAS.
GO: GO:0005886; C:plasma membrane; TAS

DR GO: GO:0016021; C:integral to membrane; TAS.
DR GO: GO:0005886; C:plasma membrane; TAS.

```

DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG-C2.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00408; IGC2; 5.
DR PROSITE; PS50835; IG-Like; 5.
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; Repeat; Signal;
KW Transmembrane; Alternative splicing.
FT SIGNAL 1 19
FT CHAIN 20 848
FT DOMAIN 20 708
FT TRANSMEM 709 729
FT DOMAIN 730 848
FT DOMAIN 20 111
FT DOMAIN 116 205
FT DOMAIN 212 301
FT DOMAIN 308 403
FT DOMAIN 406 491
FT DOMAIN 518 595
FT DOMAIN 660 727
FT DISULFID 41 96
FT DISULFID 139 189
FT DISULFID 235 287
FT DISULFID 329 385
FT DISULFID 426 479
FT CARBOHYD 222 222
FT CARBOHYD 315 315
FT CARBOHYD 347 347
FT CARBOHYD 449 449
FT CARBOHYD 478 478
FT CARBOHYD 215 215
FT CONFLICT 239 239
FT CONFLICT 480 480
FT CONFLICT 589 600
FT CONFLICT 720 721
FT CONFLICT 811 811
SQ SEQUENCE 848 AA; 93360 MW; 68D2FC0E6C1C2AD CRC64;

Query Match 5 3%; Score 122.5; DB 1; Length 848;
Beet Local Similarity 20.0%; Pred. No. 0.8;
Matches 61; Conservative 45; Mismatches 142; Indels 57; Gaps 11;

3 RG-VPFRHLLVLAVALLPATQGNKVVLGKGDVETLTCTASQKKSIOFHKNQNIKI 61
156 RGEINFKQIQVIVNVPPTIQARQINVNATANGQSTVLCVDEGFEPTMSH----- 247
62 LONQGSFLTKGPKSKLNDRASSRSRLWDQGNFLIIKNTKIEDSDTYICEV---DQKEE 117
248 -----TKDGEQIQEEDDEKYEIFSDSSQLTKIKVKNDEAEYICIAENKAGEQDAT 299
118 VQLLVFG---LTASDTHLLQGQSLTLLESPGSSPSVQCR-----SPKGNIQG- 166
300 IHLKVPFKRKITYVENQTMELBEQVTLTCEASGDPISPTIWTSTRNISSEKTLDGM 359
167 -----KTLVSQLELDQSGTWTCTVQ--NOKKVEFKIDIVLAFQKASIVYKKG 216
360 VVRSHARVSLTLKISQYTDAGEYICTASNTIGQSQSNYLEVQAPKLQGVAVYITWEG 419
217 EGVV-----FSFPLAFVYEKLTGSGELMWQARASSSSKSWITFDLKN-KEVSVKRYTODP 270
420 NQVNITCEVFAYPSA-TIS-----WPRDQLLPSSNYSNIKIYNTPSASYLEVTDPDS 470
QY 271 KQMG 275
DB 471 ENDFG 475

```

```

ANAL DROME
ID _AMAL DROME STANDARD; PRT; 333 AA.
AC P15364; OSV3A5;
DT 01-APR-1990 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Amalgam protein precursor.
GN AMA OR BG:DS00276.6 OR CG2198.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R;
RX MEDLINE=89028670; PubMed=3141062;
RT Seeger M.A., Haffley L., Kaufman T.C.;
RT "Characterization of amalgam: a member of the immunoglobulin
RT superfamily from Drosophila."
RL Cell 55:589-600(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX Celniker S.E., Pfeiffer B.D., Knafels J., Martin C.H., Mayeda C.A.,
RA Palazolo M.J.;
RA "Complete sequence of the Antennapedia complex of Drosophila."
RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutcliffe G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibegwan C.,
RA Jatalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasro P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merulio G., Milbina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Paclé J.M.,
RA Palazolo M., Peltson G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sidenkiam I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitbska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely; TISSUE=Embryo;

```

```

RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarion H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celnikier S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
CC EMBL; M23561; AAA28367.1; -.
CC EMBL; AE001572; AAD19797.1; -.
CC EMBL; AE003674; AAF54084.1; -.
CC EMBL; AY051911; AAK93335.1; -.
CC PIR; A31923; A31923.
CC FLYBASE; FBgn0000071; Ama.
CC GO; GO:0005886; C:plasma membrane; IDA.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003598; IG_c2.
CC Pfam; PF00047; Ig_3.
CC SMART; SM00408; Igc2; 2.
CC PROSITE; PS50835; IG-LIKE; 3.
KM Immunoglobulin domain; Glycoprotein; Membrane; GPI-anchor; Signal;
Repeat.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 ? AMALGAM PROTEIN.
FT PROPEP 25 333 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 139 223 IG-LIKE V-TYPE.
FT DOMAIN 230 323 IG-LIKE C2-TYPE 1.
FT DISULFID 46 117 IG-LIKE C2-TYPE 2.
FT DISULFID 161 308 PROBABLE.
FT DISULFID 251 307 PROBABLE.
FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 83 83 O -> K (IN REF. 1).
SQ SEQUENCE 333 AA; 36387 MW; F64473DE3DB25F1 CRC64;

Query Match 5.2%; Score 121.5; DB 1; Length 333;
Best Local Similarity 19.4%; Pred. No. 0.28;
Matches 56; Conservative 50; Mismatches 121; Indels 61; Gaps 9;

```

```

RESULT 42
LAMP HUMAN STANDARD; PRT; 338 AA.
ID LAMP_HUMAN
AC Q1343;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Limbic system-associated membrane protein precursor (LSAMP).
GN LSAMP OR LAMP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96235133; PubMed=8666243;
RA Pimenta A.F., Fischer I., Levitt P.;
RT "cDNA cloning and structural analysis of the human limbic-system-
RT associated membrane protein (LAMP).";
RL Gene 170:189-195(1996).
CC -1- FUNCTION: MEDIATES SELECTIVE NEURONAL GROWTH AND AXON TARGETING.
CC CONTRIBUTES TO THE GUIDANCE OF DEVELOPING AXONS AND REMODELING OF
CC MATURE CIRCUITS IN THE LIMBIC SYSTEM. ESSENTIAL FOR NORMAL GROWTH
CC OF THE HYPOCAMPAL MOSSY FIBER PROJECTION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- TISSUE SPECIFICITY: Expressed on limbic neurons and fiber tracts
CC as well as in single layers of the superior colliculus, spinal
CC chord and cerebellum.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. IGLON
CC family.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
CC EMBL; U41901; AAC50569.1; -.
CC PIR; JC4776; JC4776.
CC Genew; HGNC:6705; LSAMP.
CC MIM; 603241; -.
CC GO; GO:0007399; P:neurogenesis; TAS.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003598; IG_c2.
CC Pfam; PF00047; Ig_3.
CC SMART; SM00408; Igc2; 2.
CC PROSITE; PS50835; IG-LIKE; 3.
KM Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
Repeat; Signal; Lipoprotein.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 315 LIMBIC SYSTEM-ASSOCIATED MEMBRANE
FT PROPEP 316 338 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 132 214 IG-LIKE C2-TYPE 1.
FT DOMAIN 219 304 IG-LIKE C2-TYPE 2.
FT DISULFID 53 111 IG-LIKE C2-TYPE 3.
FT DISULFID 153 197 POTENTIAL.
FT DISULFID 239 290 POTENTIAL.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 315 315 GPI-anchor amidated asparagine
SQ SEQUENCE 338 AA; 37308 MW; 03455F286DF5D92F CRC64;

```

Query Match 5.2%; Score 121.5; DB 1; Length 338;
 Best Local Similarity 20.7%; Pred. No. 0.28;
 Matches 86; Conservative 63; Mismatches 123; Indels 143; Gaps 22;

10 LLLVQLALLPAA-----TQAKVVLGKKDPTVELTCTASQKKSIOFHKNSNQKI 61
 14 LVLALLCLLTGLVRSVDFNRGDNITVRQDTRAILRCVLEDKNS-KVAMLNRSGLIF 72
 62 LKNGSFLTKGSPKLNDR--DSRSLMDQGNFPLIKLKIEDSDTYICEVEDQK--- 116
 73 AGHD-----KMSLDPRVELEKRSLSL--EYSRIQKVVDYDGSYTCSTQTOHEBKT 121
 117 -EVLVLVFG---LANSPTHLLOGSLTLTLESPPGSSPSVQCR--SPRGKNIQGR- 168
 122 SOVYLIVQVPPKISISSDVTVNBSNVLVCMANGREPEVITWHLPTGHEFEGBEY 181
 169 LSVSOLLEDSGTWTCTVLQONQKVEFKIDIVLAFOKASSIVYKKEGBOVE--PSPFLA 226
 182 LEILGITRBOGSKVE-----KAAENVSSADVKKQVVTNYNPT 220
 227 FTVEKLTSQGEI-----WMOAERASSSKSWITFDLKNKE---V 261
 221 ITESK---SNEATYTRQASLKCEASAVPAPDFEYTRDTRINSANGL--EIKSTEGQSL 275
 262 SVKRYTQDPKLOMGKKPLHLPLPOLPOYAGSGWLTALAEKTKLQHEVNLVYMRATQ 321
 276 TVTNYTEE-----HY---GNVTCVAANKLVTN--ASLVIFR--- 307
 322 LQKNLTCEVWGPTSPKLM---LSLKLKNEKAVSKREKPVVNLNPEAGMOCLLS 373
 308 -----PGSVRGINGSISLAV-----PLWLL---ASLCLLS 336

RESULT 43
 Kilo Rat
 ID Kilo Rat STANDARD; PRT; 348 AA.
 AC 092038;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Kilon protein precursor (Kindred of Iglon).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-62.
 RX MEDLINE=99175207; PubMed=10075727;
 RA Funatsu N., Miyata S., Kumanooh H., Shigeta M., Hamada K., Endo Y.,
 RA Sokawa Y., Maekawa S.;
 RT "Characterization of a novel rat brain glycosylphosphatidylinositol-
 RT anchored protein (Kilon), a member of the Iglon cell adhesion molecule
 RT family";
 RT J. Biol. Chem. 274:8224-8230 (1999).
 RL J. Biol. Chem. 274:8224-8230 (1999).
 CC - FUNCTION: CELL-ADHESION (POTENTIAL).
 CC - SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC - TISSUE SPECIFICITY: Highly expressed in brain.
 CC - PTM: Glycosylated.
 CC - SIMILARITY: Belongs to the immunoglobulin superfamily. Iglon
 CC family.
 CC - SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC This SWISS-PROT entry is copyrighc. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL, AB017139; BAA75649.1; -.
 DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003598; Ig_C2.
 DR Pfam; PF00047; Ig_3.
 DR SMART; SM00408; IGC2; 2.
 DR PROSITE; PS00835; IG-LIKE; 3.
 KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
 KW Repeat; Signal.
 FT SIGNAL 1 31
 FT CHAIN 32 ?
 FT PROPER 32 148
 FT DOMAIN 32 128
 FT DOMAIN 133 215
 FT DOMAIN 219 307
 FT DISULFID 54 112
 FT DISULFID 154 197
 FT DISULFID 239 291
 FT CARBOHYD 67 67
 FT CARBOHYD 149 149
 FT CARBOHYD 269 269
 FT CARBOHYD 280 280
 FT CARBOHYD 288 288
 FT CARBOHYD 301 301
 SQ SEQUENCE 348 AA; 37858 MW; 37E90D1C7D2A4CAB CRC64;

Query Match 5.2%; Score 121; DB 1; Length 348;
 Best Local Similarity 23.1%; Pred. No. 0.32;
 Matches 74; Conservative 51; Mismatches 116; Indels 80; Gaps 19;

20 PAATQGNKVLGKKDPTVELTCTASQKKSIOFHKNSNQIKILNGSFLTKGP---SK 75
 35 PMAVVDNMLV--RKGDPTVLRCLYLEDGAS-KGAMLNRSGLIFAG--GDKMSVDPRVSI 89
 76 LNDADSRSLMDQGNFPLIKLKIEDSDTYICEVEDQKEVGLVGLFA----- 127
 90 LNKR-----DVSLOQNVADVTDGPFYTCSTQTOHPTPTMOYH-LTVQVPPKTYD 137
 128 -NSDTHLLOGSLTLTLESPPGSSPSVQCR--SPRGKNIQCKTLVSQLELSDGTWTC 184
 138 ISNDMTINEGTVTLTCLATKPEPAISWRHISPAKFEQVLYDIGTRDQAGEVREC 197
 185 TVLONQKVEF---KIDIVV---LAFQKASSI-----VYKKEGQVEFSPPLATTV 229
 198 SA---ENNVSPDVKKVAVVNFAPFTIOEIKSGTVTPGRSLIRCEGAVP---PPAPE- 250
 220 EKLKSGSLMWQARRA---SSSKSWITFDLKNKEV-SVKRYTOD-----PKL-QMG 275
 251 -----WYKGEGRFLNQGQGIIONFSRSLITVNTVOEHGNTVCAANKLGTIN 301
 276 KKLPLHLPLPOLPOYAGSGN 296
 302 ASLPLN---PPTAQYGTGS 319

RESULT 44
 SCP1 Rat
 ID SCP1 Rat STANDARD; PRT; 997 AA.
 AC 003410;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 15-JUN-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Synaptonemal complex protein 1 (SCP-1 protein).
 GN SYCP1 OR SCP1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9309684; PubMed=1464329;
 RA Meuwissen R.L.J., Offenberg H.H., Dietrich A.J., Riesewijk A.,
 RA Iersel M., Heyting C.;
 RT "A coiled-coil related protein specific for synapsed regions of
 RT meiotic prophase chromosomes";

```

RL EMBO J. 11:5091-5100(1992).
CC -1- FUNCTION: Major component of the transverse filaments of
CC synaptonemal complexes (SCs), formed between homologous
CC chromosomes during meiotic prophase.
CC -1- SUBCELLULAR LOCATION: Nuclear. In tripartite segments of
CC synaptonemal complexes, between lateral elements in the nucleus.
CC Found only where the chromosome cores are synapsed. Its N-terminus
CC is found towards the centre of the synaptonemal complex while the
CC C-terminus extends well into the lateral domain of the
CC synaptonemal complex.
CC -1- TISSUE SPECIFICITY: Testis.
CC -1- DEVELOPMENTAL STAGE: Expressed exclusively in meiotic prophase
CC cells.
CC -1- DOMAIN: Consists of an alpha-helical stretch of 700 AA residues,
CC flanked by N- and C-terminal globular domains. The C-terminal
CC domain has DNA-binding capacity.
CC -1- CAUTION: This is a conceptual translation; a frameshift was
CC corrected in position 6 to maximize the similarity with the
CC other species SyCP1 sequences.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X67805; CA48006.1; ALT_FRAME.
DR InterPro: IPR008827; SCP-1.
DR Pfam: PF05483; SCP-1;
KM Nuclear protein; Meiosis; Cell division; Phosphorylation;
FT DNA-binding; Coiled coil.
FT DOMAIN 108 819 COILED COIL (POTENTIAL).
FT DOMAIN 118 121 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 701 704 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 902 905 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 982 990 ARG/LYS-RICH (BASIC).
SQ SEQUENCE 997 AA; 116511 MW; 229D59823FD684BE CRC64;

Query Match 5.2%; Score 121; DB 1; Length 997;
Best Local Similarity 19.8%; Pred. No. 1.2; Indels 118; Gaps 24;
Matches 96; Conservative 95; Mismatches 175;

QY 9 HLLVLTQ-----ALPPATQGNKVLGKKGDVELTCTASQKSIQF-----WKNNOIK 60
DB 374 HSLVTELKATCTLEELRTQGLNENEDQKIMELQKSSLEEMKFRKNNEVE 433
QY 61 ILNGSGFLTKGPKSLNDR-----ADSRSLMDQGNFPLIITKNKIED----- 103
DB 434 -LEELKTLIAEDQKLLDEKQVEKLAELQCKEDELTLQTRKEKIHDELVQVTVTKTS 492
QY 104 SDTYCEVEDQKEVQ-----LVFGLTANSDFHLLQGGSLT-----LTLESPPSSPSQOC 155
DB 493 ESHYKQVEEMTELEKEKLNIELTANSDDLLENKKLVQDASDMVELKKHODIINC 552
QY 156 RSPRKNQGGKTLVSQLELQDS-GTWTCVLOKQKVEFKIDIVLAFQKASSIVY-- 212
DB 553 KKQBERMLKQIETLEKEMNLDELSEVRKEIIOGGDEVKCKLD--KSEENASIEVEV 609
QY 213 -KKEGE-----QVEFSPPLAFVTEKLTGSGELMWMQERASSSKSWTFDIK- 257
DB 610 LKKEKQMKILEKCNLKKQIENK--SKNIEELHQENKA-LKKKSSAKENQLNAVEIKV 665
QY 258 -----NKEYSVRKRVNDPRLQMGKGLPLHLTLPOLAPYAGSGN 296
DB 666 NKLIELASTQKPEEMINNYQKEITEIKKISE-RL-LGEVEKAYATDEAVKL--OKE 720
QY 297 LTLALEATGKL-----HOEVALVVMRATQ-LQNLTCCEVWGPTSPKMLSTLKEN 346
DB 721 IDLRQGHIAEVALMEKHQYDKIYVERDSELGLYNKRQD--QSSAYALLETLSN 777
QY 347 -----KEAKVSKREKRVVVLNPEAGMOCILSDSGVLLBSNRYKLP--TW----- 390

```

```

DB 778 IRNELVSLKQLVEKEKEKELKMEQEN---TAILTDKKDKIQASLSEAFSRTS 834
QY 391 -STP 393
DB 835 KTTIP 838

RESULT 45
PIGR_RABIT
ID PIGR_RABIT STANDARD; PRT; 773 AA.
AC P01832;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Polymeric-immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR)
DE [Contains: Secretory component].
GN PIGR.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84142246; PubMed=6322002;
RA Mostov K.E., Friedlander M., Blobel G.;
RT "The receptor for transepithelial transport of IGA and IGM contains
RT multiple immunoglobulin-like domains".
RL Nature 308:37-43(1984).
RN [2]
RP SEQUENCE OF 87-114 AND 410-428.
RX MEDLINE=88228032; PubMed=313339;
RA Frutiger S., Hughes G.J., Hanly W.C., Jaton J.-C.;
RT "Rabbit secretory components of different allotypes vary in their
RT carbohydrate content and their sites of N-linked glycosylation.";
RL J. Biol. Chem. 263:8120-8125(1988).
CC -1- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
CC TRANSMEMBRANE SEGMENT.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Also secreted.
CC -1- POLYMORPHISM: The sequence shown is that of allotype T62.
CC -1- SIMILARITY: Contains 5 immunoglobulin-like V-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X00412; CA425118.1; -.
DR PIR: A02111; ORRBG.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003599; IG.
DR Pfam: PF00047; Ig; 5.
DR SMART: SM00409; Ig; 5.
DR PROSITE: PS50835; IG_LIKE_3.
KW Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal;
KW Polymorphism.
FT SIGNAL 1 18
FT CHAIN 19 773 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
FT CHAIN 19 615 SECRETORY COMPONENT.
FT DOMAIN 19 647 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 648 670 POTENTIAL.
FT DOMAIN 671 773 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 131 IG-LIKE V-TYPE 1.
FT DOMAIN 138 232 IG-LIKE V-TYPE 2.
FT DOMAIN 233 340 IG-LIKE V-TYPE 3.
FT DOMAIN 352 455 IG-LIKE V-TYPE 4.

```

```

FT DOMAIN 461 557 IG-LIKE V-TYPE 5.
FT DISULFID 46 115 POTENTIAL.
FT DISULFID 155 225 POTENTIAL.
FT DISULFID 260 324 POTENTIAL.
FT DISULFID 369 438 POTENTIAL.
FT DISULFID 478 538 POTENTIAL.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (IN ALLOTYPES T61).
FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (IN ALLOTYPES T62).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (IN ALLOTYPES T63).
FT VARIANT 88 88 K -> N (IN ALLOTYPES T61).
FT VARIANT 94 94 D -> E (IN ALLOTYPES T61).
FT VARIANT 101 108 TYDOLTON -> YLNRLSOS (IN ALLOTYPES T61).
FT VARIANT 110 110 S -> T (IN ALLOTYPES T63).
SQ SEQUENCE 773 AA, 83866 MW, DP2C44D2F1193C65 CR664;

Query Match 5.2%; Score 120.5; DB 1; Length 773;
Best Local Similarity 20.3%; Pred. No. 0.97;
Matches 108; Conservative 75; Mismatches 199; Indels 149; Gaps 26;

QY 5 VFFRHL-----VLQALLPAATQGN-----KVLGKKGTVELTCTASQKSIQ 50
DB 210 VIKHLQLNDAGQYQSGSDPTAEQNDRLRLTGLLYGNIGSGVTFECALDSBDANA 269
QY 51 FPMKSNQIK-----ILNGSGFLTKGPKSLNDRASRLM---DQGNPLIKLIKIED 103
DB 270 V--ASLRVGRGNVVIDSGTI-----DPAFGRLLFTKAENGHSVIAIRKED 318
QY 104 SDTYICEVDEKQEVOLVFGLTANSDFHLQ---GQSILTLTLESP---PGSSPSVQC 155
DB 319 TGNVLCGVQSN-----GQSGDPTQLRLQFVNEEIDVGRSPVLKGFQGSVTRC 369
QY 156 -RSPP-----GKNI-----QGKITSV---SQLE 175
DB 370 PYNPKRSDSHQLYLMEGSGQTRHLVDSGEGVLQDYTRLALFEPGNGTSSVYLNQLT 429
QY 176 LDPSGTWTCTVLQNKQKVEFKIDIVLAFQKASSI--VYKKEGEOVEFS--PPLAFVYEX 231
DB 430 AEDGFGYWC-VSDDDESILTSYKQLQVDESPFTIDKFAVQGEPEITCHFPCKY---- 484
QY 232 LTSGSELMQAERASSSSKSMITF-DLKNKEVSVKRVTOPDKLQMGKULPLHLTPQALPQ 290
DB 485 -----FSEKWKYCKMNDHGCDLPTKLSSSDGLVKCNMNLVLTLDVSVED 531
QY 291 YG---SGNLTALAEKTKLHGVNLVVRATLOLKNLTCEWGPSTSKMLSLKLENK 347
DB 532 DEGWYWC-----AKDGHFEFEVAAVAVELTEPAKVAEPAKVVDPAKAPAPAEEX 584
QY 348 EAKVSRKREKPVWVLANPEAGMGCCLDSQVLLSENIKVLPTWSTPVHPRASALDAPPTG 407
DB 585 ----AKARCPV---PRRQWYPL-----SRKLRT-SCP-EPRLIAEEVAVOS 622
QY 408 SALPPQTSALPDPPASALPALAVISFLGLGI-----GVACYLAKTR 453
DB 623 ADDPASGSAVSADASSGSGSAKVLSTVPLGLVLAAGMAVAIAAR 673

RESULT 46
NCAL CHICK STANDARD; PRT; 1091 AA.
AC P13590; Q90918; Q90919;
DT 01-JAN-1990 (rel. 13, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE Neural cell adhesion molecule 1, 180 kDa isoform precursor (N-CAM
DE 180).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxId=9031;
RN [1]
RP SEQUENCE OF 1-175 FROM N.A., AND PARTIAL SEQUENCE.

```

```

RX MEDLINE=87206190; PubMed=3576199;
RA Cunningham B.A., Hemperly J.J., Murray B.A., Prediger E.A.,
RA Brackenbury R., Edelman G.M.;
RT "Neural cell adhesion molecule: structure, immunoglobulin-like
RT domain, cell surface modulation, and alternative RNA splicing.";
RL Science 236:799-806(1987).
RN [2]
RP SEQUENCE OF 128-1091 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=86206089; PubMed=3458261;
RA Hemperly J.J., Murray B.A., Edelman G.M., Cunningham B.A.;
RT "Sequence of a cDNA clone encoding the polysialic acid-rich and
RT cytoplasmic domains of the neural cell adhesion molecule N-CAM.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:3037-3041(1986).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM D).
RX MEDLINE=87092340; PubMed=3467341;
RA Hemperly J.J., Edelman G.M., Cunningham B.A.;
RT "CDNA clones of the neural cell adhesion molecule (N-CAM) lacking a
RT membrane-spanning region consistent with evidence for membrane
RT attachment via a phosphatidylinositol intermediate.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:9822-9826(1986).
RN [4]
RP SEQUENCE OF 810-1069 FROM N.A.
RX MEDLINE=87033934; PubMed=3771645;
RA Murray B.A., Owens G.C., Prediger E.A., Crossin K.L.,
RA Cunningham B.A., Edelman G.M.;
RT "Cell surface modulation of the neural cell adhesion molecule
RT resulting from alternative mRNA splicing in a tissue-specific
RT developmental sequence.";
RL J. Cell Biol. 103:1431-1439(1986).
RN [5]
RP SEQUENCE OF 1-17 FROM N.A.
RX MEDLINE=93122797; PubMed=1478668;
RA Colwell G., Li B., Forrest D., Brackenbury R.;
RT "Conserved regulatory elements in the promoter region of the N-CAM
RT gene.";
RL Genomics 14:875-882(1992).
RN [6]
RP SEQUENCE OF 1-17 FROM N.A.
RC STRAIN=white leghorn; TISSUE=Erythrocyte;
RA Saecker M., Covault J.;
RT Submitted (AUG-1993) to the EMBL/Genbank/DBJ databases.
CC -! FUNCTION: This protein is a cell adhesion molecule involved in
CC neuron-neuron adhesion, neurite fasciculation, outgrowth of
CC neurites, etc.
CC -! ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=A; Synonyms=N-CAM 180;
CC IsoId=P13590-1; Sequence=Displayed;
CC Name=B; Synonyms=N-CAM 140; VSP_002585;
CC IsoId=P13590-2; Sequence=VSP_002585;
CC Name=C;
CC IsoId=P13590-3; Sequence=VSP_002586;
CC Name=D;
CC IsoId=P13590-4; Sequence=VSP_002583, VSP_002584;
CC -! SIMILARITY: Contains 2 fibronectin type III domains.
CC -! SIMILARITY: Contains 2 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15861; AAB59958.1; JOINED.
DR EMBL; M15860; AAB59958.1; JOINED.
DR EMBL; M15922; AAB59958.1; JOINED.
DR EMBL; M15923; AAB59958.1; JOINED.
DR EMBL; M15924; AAB59958.1; JOINED.
DR EMBL; M21178; AAB59958.1; JOINED.
DR EMBL; M21179; AAB59958.1; JOINED.

```



```

RC TISSUE=Fat body;
RX MEDLINE=97166603; PubMed=9014323;
RA Sundermeyer K., Hendricks J.K., Prasad S.V., Wells M.A.;
RT "The precursor protein of the structural apolipoproteins of
  lipophorin: cDNA and deduced amino acid sequence.";
RT Insect Biochem. Mol. Biol. 26:735-738(1996).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; SECRETED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.trib-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL, U57651; AAB53254.1; -.
DR PIR, T18358; T18358.
DR InterPro: IPR001747; Lipid_transprt_N.
DR InterPro: IPR001846; VWF_D.
DR Pfam: PF01347; Vitellogenin_N; 1.
DR Pfam: PF00094; vwd; 1.
DR SMART, SM00638; LPD_N; 1.
DR SMART, SM00216; VWD; 1.
KW Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 ? APOLIPOPHORIN-11.
FT CHAIN ? 3305 APOLIPOPHORIN-1.
SQ SEQUENCE 3305 AA; 366941 MW; B27D611410285FD7 CRC64;

Query Match
Best Local Similarity 5.2%; Score 120.5; DB 1; Length 3305;
Matches 97; Conservative 66; Mismatches 176; Indels 93; Gaps 21;

QY 26 NKRYLGGKQDTVELTCTASQKKS-----IQFMKNSNQIKIGNSQFLTKGSPKXLDRA 80
DB 1852 NSVVVDADQRYVKIDSSIVLSKAHPVLDIQYHSPSSDKIRRLYLQSSLSSTQGLEVKV 1911
QY 81 DSRRLMDQGNFPLIINKIKIEDSDTYICEVEDQKEVQLVFGLTANSDDTLLOGQSIT 140
DB 1912 DN-----INDICLD---AVSEANVQKDN---AFKVANAKELGWMKNGID 1951
QY 141 LTLSPGSSPSPVOCRSPPR-GKNIQGGKTLVSQLELQDS--GTWTCVLONQKKEVP 196
DB 1952 IS-SKDSGSGKRLPHATDNKNVLSGSTSFISKQEGQTTIEGSGSVKVEEQSANPK 2010
QY 197 IDIVVIAFOKASSIYVKKEGEO-VFSPPLAFTVEKLTSGGELMWOAERASSSKSWIT-F 254
DB 2011 Y-----IRTVFTDSNKEGVETFFVVAL-----GERSYVAESRVITNY 2046
QY 255 DLKAKKESVAVKVTQDPKLOMGKKLPVHLTLPOALPOYAGSGMLTLAEKTKGLHQEVNL 314
DB 2047 EYKNSYVVCSEKKQCAHAEIQSK--IDMSTPGMIVNINAG-LDL---RKLGV----- 2093
QY 315 VVMRAATOLQKMLTCVWGPTSPKMLSLKLENKEAKVSRKRPVAVLN-PEAGMMQCLLS 373
DB 2094 ----APBGLQMRDEVSRRPRFTLDLHT-NKE---DRKHLLAAVYTPBEGHY----A 2140
QY 374 DSGQVVLBSNIKVLPFTWSPVHPRASALPAPRTGSALPD-----DOTASALPDPRA 425
DB 2141 SGVTVRLPSPRVVALE--YTLTHTPTSQDLPFPFKGEACLDLDKNRPGHKTSARFLVDYNS 2198
QY 426 SALPALAVISF 437
DB 2199 GSEDKAVAEIGF 2210

RESULT 48
PECT1_PIG STANDARD; PRT; 740 AA.
AC 095242;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

```

```

DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Platelet endothelial cell adhesion molecule precursor (PECAM-1)
DE (CD31 antigen).
GN PECAM1.
OS Sus scrofa (Pig).
OC Euxaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxId=96823;
RN [1]
RP SEQUENCE FROM N.A.
RA Nasu K.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE EXPRESSED ON
  PLATELETS AND AT ENDOTHELIAL CELL INTERCELLULAR JUNCTIONS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.trib-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL, X98505; CA67129.1; -.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003599; Ig.
DR Pfam, PF00047; Ig; 4.
DR SMART, SM00409; Ig; 3.
DR PROSITE, PS50835; IG_LIKE; 4.
KW Immunoglobulin domain; Glycoprotein; Cell adhesion; Transmembrane;
KW Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 740
FT DOMAIN 28 602 BY SIMILARITY.
FT TRANSMEM 603 621 PLATELET ENDOTHELIAL CELL ADHESION
FT DOMAIN 622 740 MOLECULE.
FT DOMAIN 35 126 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 145 223 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 236 315 IG-LIKE C2-TYPE 1.
FT DOMAIN 328 404 IG-LIKE C2-TYPE 2.
FT DOMAIN 425 494 IG-LIKE C2-TYPE 3.
FT DOMAIN 500 592 IG-LIKE C2-TYPE 4.
FT DISULFID 57 109 IG-LIKE C2-TYPE 5.
FT DISULFID 152 206 IG-LIKE C2-TYPE 6.
FT DISULFID 256 304 POTENTIAL.
FT DISULFID 347 387 POTENTIAL.
FT DISULFID 432 477 POTENTIAL.
FT DISULFID 524 573 POTENTIAL.
FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 84 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 436 436 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 552 552 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 740 AA; 82378 MW; F312DC624B4A217 CRC64;

Query Match
Best Local Similarity 5.1%; Score 118.5; DB 1; Length 740;
Matches 75; Conservative 58; Mismatches 116; Indels 99; Gaps 19;

QY 101 IEEDDTYICEVEDQK-BEVQLLVFGLT-----ANSDFHLQGGSLITLLESPRESSP 151
DB 296 VEHNSNTCKVEARISRVSSIMVNTLFSRPLKKSATRLDGGESIRLMC-SIPGAP 354
QY 152 SVQKSPRGKNIQGGKTLVSQLEL-----QDSGTWTC-----YLQNGKQVFKKIDIV 200

```

```

Db 355 EAMF-----TIQKGMMLDQDNLTKVASERDSGTTCVAGIGKVKRSNEVOIJCEN 408
Oy 201 V---LAFQKASSIYKKEGEVFEFS-----FPLAFTVEKLTGSGELMWOAERRASSKS 250
Db 409 LSKPSIFDSSGSEVKK--GQITIEVCSQINGSTPSISTY--QLIKSSDL--ASQNVSSNEP 462
Oy 251 WITFDLKNKEVSVKRVTDQPKLQMGKULPLHLTLPLQALPOYAGSGNLTLEAKTGKLIHQ 310
Db 463 AVFKONPKRDVEYQCIADN-----CHSHAGMPSKVLRVKVIAPVE 502
Oy 311 EVNLVVMATQLOQRK---LTCEV---WGPTSPKMLSLKLNKRAKSKREKP----- 357
Db 503 EVKLISLSEVESGQALVLQCSYKSGSGPTTKFY-----KX---KENKPFQVTL 551
Oy 358 -----VM-----VLNPEAGMOCCLSD---SGOVLESNIKVLPTWSTP 393
Db 552 NDTQAIWKPKRASKQEGQYICLASNRATPSKNFLQSNITLAVRYLAP 599

RESULT 49
LAMP_RAT ID LAMP_RAT STANDARD; PRT; 338 AA.
AC Q62813;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Limbic system-associated membrane protein precursor (LSAMP).
GN LSAMP OR LAMP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N. A., AND SEQUENCE OF 29-49.
RC TISSUE=Hippocampus;
RX MEDLINE=95374785; PubMed=7646886;
RA Pimenta A.F.; Zhukareva V.; Barbe M.F.; Reinoso B.S.; Grimley C.;
RA Henzel W.; Fischer I.; Levitt P.;
RT "The limbic system-associated membrane protein is an Ig superfamily
RT member that mediates selective neuronal growth and axon targeting.";
RT Neuron 15:287-297(1995).
RL
CC -I- FUNCTION: MEDIATES SELECTIVE NEURONAL GROWTH AND AXON TARGETING.
CC CONTRIBUTES TO THE GUIDANCE OF DEVELOPING AXONS AND REMODELING OF
CC MATURE CIRCUITS IN THE LIMBIC SYSTEM. ESSENTIAL FOR NORMAL GROWTH
CC OF THE HIPPOCAMPAL MOSSY FIBER PROJECTION.
CC -I- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -I- TISSUE SPECIFICITY: Expressed mostly by neurons comprising limbic-
CC associated cortical and subcortical regions that function in
CC cognition, emotion, memory, and learning.
CC -I- DEVELOPMENTAL STAGE: FIRST DETECTED AT E15-16, AT STAGE E20 IT IS
CC DETECTED IN PRESUMPTIVE CORTEX, MEDIAL LIMBIC AREAS OF THE
CC THALAMUS AND HYPOTHALAMUS. IN THE ADULT, IT IS FOUND IN
CC HYPOTHALAMUS, PERIRHINAL CORTEX, AMYGDALA AND MEDIAL THALAMIC
CC REGION.
CC -I- SIMILARITY: Belongs to the immunoglobulin superfamily. IGLON
CC family.
CC -I- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U31554; AA86120.1; .
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_C2.
CC Pfam; PF00047; Ig_3.
CC SMART; SM00408; IGC2; 2.
CC PROSITE; PS50835; Ig_LIKE; 3.
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;

```

```

KW Repeat; signal; Lipoprotein.
FT SIGNAL 1 28
FT CHAIN 29 315
FT PROPEP 316 338
FT DOMAIN 29 122
FT DOMAIN 132 214
FT DOMAIN 219 304
FT DISULFID 53 111
FT DISULFID 153 197
FT DISULFID 239 290
FT CARBOHYD 40 40
FT CARBOHYD 66 66
FT CARBOHYD 136 136
FT CARBOHYD 148 148
FT CARBOHYD 279 279
FT CARBOHYD 287 287
FT CARBOHYD 300 300
FT CARBOHYD 315 315
FT LIPID 315 315
SQ SEQUENCE 338 AA; 37324 MW; 0B76AFDD68A39BB6 CRC64;

Query Match 5.1%; Score 117.5; DB 1; Length 338;
Best local similarity 20.5%; Pred. No. 0.53;
Matches 85; Conservative 64; Mismatches 123; Indels 143; Gaps 22;

Oy 10 LLVLVQLALPLPA-----TQGNKVLGKGGDTVELTCTAQGKKSIOFHKNSNQIKI 61
Db 14 LVLRRLCLPLPTGLPVRVDFNRGTNITVROGDTALRLCVEDKNS-KVAMLRSGIIF 72
Oy 62 LGNGSFLTKGPKSLNDRA--DSRRSLMDQGNPLIINKIEDSDTYICEVEDQKE--- 116
Db 73 AGHD-----KMSLDPRVLEKRNHL-----EVSRLQKVDYDEGYSVQGHPRKT 121
Oy 117 -EVQLVFG---LTNSDTHLIQGSGLTTLSPSPSSVSQCR--SPGKNIQSGKT- 168
Db 122 SOVYLLIQVPPKISNISSDVTNVEGSNVTLVCNAGRPREVITWRHLTPLGREFEGEEY 181
Oy 169 LVSQLELDQSGTWTCTVQLNQKVERKIDIVLAFQKASSIYKKEGEVFE--FSFPLA 226
Db 182 LEIIGTIREOSGKEC-----KANEVSSADAVKQKVTNVNPP 220
Oy 227 FVEKLTGSGEL-----WMOAERRASSKSWITPLKNKE---V 261
Db 221 ITSEK---SNEATTGQASLKCEASAVPAPDFEYRDRDTRINSANGI--EIKSTEGGSSL 275
Oy 262 SVKRVTDQPKLQMGKULPLHLTLPLQALPOYAGSGNLTLEAKTGKHQEVNLVVMRATQ 321
Db 276 TVTNVTEE-----HY-----GNVTCVAANKLGVTN--ASLVLFR--- 307
Oy 322 LQKNLTGEVWGPTSPKLM---LSIKLENKRAKSKREKPVWVNLNPEAGMOCCLIS 373
Db 308 -----PGSVGINGSISLAV-----PLMLL---AASLFCILS 336

RESULT 50
NCAL_RAT ID NCAL_RAT STANDARD; PRT; 858 AA.
AC P13596;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neural cell1 adhesion molecule 1, 140 kDa isoform precursor (N-CAM 140)
DE (NCAM-140).
GN NCAM1 OR NCAM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N. A.
RC TISSUE=Brain;

```

```

RX MEDLINE=88059265; PubMed=3680385;
RA Small S.J., Shull G.E., Santoni M.-J., Akesson R.;
RT "Identification of a cDNA clone that contains the complete coding
RL sequence for a 140-kD rat NCAM polypeptide.";
RN J. Cell Biol. 105:2335-2345(1987).
[2]
RP SEQUENCE OF 355-364 FROM N.A.
RX MEDLINE=90166485; PubMed=2483093;
RA Small S.J., Haines S.L., Akesson R.A.;
RT "Polypeptide variation in an N-CAM extracellular immunoglobulin-like
RL fold is developmentally regulated through alternative splicing.";
RN Neuron 1:1007-1017(1988).
CC -1- FUNCTION: This protein is a cell adhesion molecule involved in
neuron-neuron adhesion, neurite fasciculation, outgrowth of
neutres, etc.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=1;
Comment=a number of isoforms are produced;
Name=1;
CC -1- IsoId=P1356-1; Sequence=displayed;
CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; X06564; CAA29809.1; -
CC EMBL; M32611; AAA41679.1; -
CC PIR; S00846; IJRTNC.
CC PDB; 1BPF; 27-OCT-00.
CC InterPro; IPR008957; FN_III-like.
CC InterPro; IPR003961; FN_III.
CC InterPro; IPR007110; Ig_IIIc.
CC InterPro; IPR003598; Ig_C2.
CC Pfam; PF00041; fn3; 2.
CC Pfam; PF00047; Ig; 5.
CC SMART; SM00060; FN3; 2.
CC SMART; SM00408; IgC2; 5.
CC PROSITE; PS50835; Ig_Like; 5.
CC Cell adhesion; Glycoprotein; Transmembrane; Repeat;
CC Immunoglobulin domain; Alternative splicing; Signal; Heparin-binding;
CC 3D-structure.
CC SIGNAL 1 19
CC CHAIN 20 858
CC FT DOMAIN 20 721
CC FT TRANSMEM 722 739
CC FT DOMAIN 740 858
CC FT DOMAIN 20 111
CC FT DOMAIN 116 205
CC FT DOMAIN 212 302
CC FT DOMAIN 309 414
CC FT DOMAIN 417 502
CC FT DOMAIN 514 615
CC FT DOMAIN 616 712
CC FT DOMAIN 712 732
CC FT DOMAIN 152 156
CC FT DOMAIN 161 165
CC FT DISULFID 41 96
CC FT DISULFID 139 189
CC FT DISULFID 235 288
CC FT DISULFID 330 396
CC FT DISULFID 437 490
CC FT CARBOHYD 222 222
CC FT CARBOHYD 316 316
CC FT CARBOHYD 348 348
CC FT CARBOHYD 434 434
CC FT CARBOHYD 460 460

```

```

FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 858 AA; 94658 MW; EA1A06A4EA0550F6 CMC64;
Query Match
Best Local Similarity 5.1%; Score 117.5; DB 1; Length 858;
Matches 79; Conservative 52; Mismatches 143; Indels 139; Gaps 17;
QY 3 RG-VFRRILLVLQALALPAATQGNKVVLGKKDPTVELTCAQSKKSIOFFHM-KNSNQIK 60
DB 196 RGEINFKIOIVIVNVPVQAROSTIVATYANLQGSVTLVCADGFPPEFTMSMTXDGPIE 255
QY 61 ILNGGSPFLTGKPSLRADSRSLMDQGNFPLIKLKIETSDTYICEVE---DQKE 116
DB 256 -----NEEDDEKHFISDDSELTIRVNDKDAEYCIENKAGEBDA 299
QY 117 EVQLLVFG---LRANSPTHLIQGSLTTLTLESPGSSPVQCSPPKQNIQGGKTLSSV 172
DB 300 SIHLKVFAPKITYVENQTAMELEEQVTLTCEASDGPISITWRT-----STR 347
QY 173 QLELDSDGTWCTVYQNGKQKVEFKIDIVLAFQKASIVYKKEGQVFEFPLAFTVKL 232
DB 348 NISSEKASWT---RPEKQETLGGHWVRSHARVSLTLK---SIQY----- 388
QY 233 TGSGETLMQARERASSSKSMITFDLKNKEVSVKRVTDPKLQMGKKLPHLTLPOALPOYA 292
DB 389 TDAGVYICTAINTIGQDS-----QSMYLEVQYAPAKLG-----PVAV--YT 427
QY 293 GSGNITTLAEAKTKLHQEVNLVVMRATQKNTLCEVWGPTSPKMLSLKENKAVS 352
DB 428 WEGN-----QVNITCEVFAYPS-----ATIS 448
QY 333 KREKPVWLNPEAGMOCCLSDGOVLLESNIKVLPTSTVHPRASALPRP 405
DB 449 -----WPRD-----GQLLPSSVSNYNIKIYNT---PSASYLEVTP 479
RESULT 51
CAML BRARE
ID CAML_BRARE STANDARD; PRT; 1197 AA.
AC Q90478;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neutral cell adhesion molecule LI.1 (N-CAM LI.1) (Fragment).
GN NADL1.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7955;
RN (1)
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RX TISSUE=embryo;
RX MEDLINE=96155762; PubMed=8568941;
RA Tongiorgi E., Bernhardt R.R., Schachner M.;
RT "Zebrafish neurons express two LI-related molecules during early
axogenesis.";
RL J. Neurosci. Res. 42:547-561(1995).
CC -1- FUNCTION: Cell adhesion molecule with an important role in the
development of the nervous system. Involved in neuron-neuron
adhesion, neurite fasciculation, outgrowth of neurites, etc. Binds
to axonin on neurons (by similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in postmitotic neurones in 16-36
hour embryos, including those in the brain, cranial ganglia and
otic and olfactory placodes, and in all classes of spinal
neurons.
CC -1- DEVELOPMENTAL STAGE: Onset of expression correlates with the
initiation of axonogenesis in 16-36 hour embryos.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily.
CC LI/neurofascin/NCAM family.
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 5 fibronectin type III domains.

```

```

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X89204; CAA61490.1; -.
DR PIR: T30581; T30581.
DR HSSP: P20241; 1C6B.
DR ZFIN: ZDB-GENE-960526-512; nad11.1.
DR InterPro: IPR008957; FN_III-like.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003962; FnIII subd.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig_C2.
DR Pfam: PF00041; fn3; 5.
DR Pfam: PF00047; fn3; 6.
DR PRINTS: PR00014; ENTPEP11.
DR SMART: SM00060; FN3; 5.
DR SMART: SM00408; IGC2; 4.
DR PROSITE: PS50835; IG_Like; 6.
DR Neurogenesis; Cell adhesion; Developmental protein; Glycoprotein;
DR Transmembrane; Repeat; Immunoglobulin domain.
KW NON TER
FT DOMAIN 1 1054 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1055 1075 POTENTIAL.
FT DOMAIN 1076 1197 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1 58 IG-LIKE C2-TYPE 1.
FT DOMAIN 165 160 IG-LIKE C2-TYPE 3.
FT DOMAIN 263 355 IG-LIKE C2-TYPE 4.
FT DOMAIN 360 442 IG-LIKE C2-TYPE 5.
FT DOMAIN 451 541 IG-LIKE C2-TYPE 6.
FT DOMAIN 546 638 FIBRONECTIN TYPE-III 1.
FT DOMAIN 645 739 FIBRONECTIN TYPE-III 2.
FT DOMAIN 744 849 FIBRONECTIN TYPE-III 3.
FT DOMAIN 850 948 FIBRONECTIN TYPE-III 4.
FT DOMAIN 952 1029 FIBRONECTIN TYPE-III 5.
FT DISULFID 92 143 BY SIMILARITY.
FT DISULFID 199 247 BY SIMILARITY.
FT DISULFID 289 339 BY SIMILARITY.
FT DISULFID 383 432 BY SIMILARITY.
FT DISULFID 472 525 BY SIMILARITY.
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 449 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 708 708 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 959 959 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1002 1002 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1027 1027 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1197 AA; 132860 MW; 7CB1505EEF7C7B28 CRC64;
-----
Query Match 5.1%; Score 117.5; DB 1; Length 1197;
Best Local Similarity 18.0%; Pred. No. 2.7;
Matches 83; Conservative 84; Mismatches 172; Indels 123; Gaps 19;
-----
OY 9 HLLVLQLALPAAQGNKVVLSGGKGVLELTCTASQKKSIOFHMKNNOIKILGNQGSF 68
DB 259 HVTIVTEAA--PYWTRSPBEHLVAPGETVRLDCKADGIPAPITW-SINGVPVSGTD--- 312
OY 69 LTKGSKLNDRAISRSLMDQGNFPLIKNLKIEISDTYICEVEDQKREYQVLVFGULAN 128
DB 313 -----VDPKRRV---SSGKLLISNVEFSDTAVYQCAVAKGSIIL-----N 351

```

```

OY 129 SDTHLQGSILTLT-----LESPPGSSPVQCRS-----PRGNIGCGK 167
DB 352 THYHVELPQIILLPDLRLOVATAGQVMDLCTFGSPPLKRIHEILDSIPALSNAKISQ 411
OY 168 T---LSVQLELQDSCTWTCTYLQNKQKVEFKIDV-----VLAFOQASSIVY 212
DB 412 TTNGSLKISVNSEDSNRKTCVSETKKISADAEVLRKTVGPDPNHLVIRSGDILH 471
OY 213 KKEGEQVEFSFPLAFITVEKTLTSGELWQALERASSKSWITPDKNKEVSKVRYTOPDK 272
DB 472 CK-----YTDHNLKSPYQVNMNDGKHITNS--TSNDKHEIFESGLKVLVDQM 517
OY 273 Q-WCKKLPLHLTLPLQALPOYAGSGNLTALBAATKTL-----HQEVNLVVMRATQLQ 323
DB 518 EDWG-----ISCEVSTLTDSDTASGVITVQDKPDPQSLKL-----SEKME 559
OY 324 KNLTCFVWGPVTL---SPKMLSLKLEKKEKAKVSKREKPVWLNPPEAGMQLCLSDSQVL 379
DB 560 RSVYIS-WMPSVENNSVTEVEIEMEGE-----TPDEQWQKYSRVSQDID 605
OY 380 LESNIXVLPTWSPVPHRASALPAPPTGSLPDPQTASALPD 421
DB 606 SMRSICSYSKYHROIARVNSIGTSAPTESL-SYSTPAAPD 646
-----
RESULT 52
C166_CARAV STANDARD; PRT; 555 AA.
ID C166_CARAV
AC Q90304;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE CD166 antigen homolog precursor (Neuroilin) (DW-GRASP homolog).
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Laessing U.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 30-552 FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Retina;
RX MEDLINE=94299040; PubMed=8026643;
RA Laessing U., Giorzano S., Stecher B., Lottspeich F., Stuermer C.A.O.;
RT "Molecular characterization of fish neuroilin: a growth-associated
RT cell surface protein and member of the immunoglobulin superfamily in
RT the fish retinorectal system with similarities to chick protein
RT DM-GRASP/SC-1/BEN.";
RL Differentiation 56:21-29(1994).
CC -1- FUNCTION: CELL ADHESION MOLECULE. INVOLVED IN NEURITE EXTENSION BY
CC NEURONS VIA HETEROPHILIC AND HOMOPHILIC INTERACTIONS.
CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
CC -1- TISSUE SPECIFICITY: PRESENT ON ALL RETINAL GANGLION CELLS (RGCs)
CC AND THEIR AXONS (IN EMBRYO). ABSENT FROM MATURE AXONS ALONG MOST
CC OF THEIR LENGTH, BUT IS PRESENT ON NEW AND GROWING AXONS DERIVED
CC FROM THE RGS AT THE RETINAL MARGIN. REMAINS ON ADULT RGS ONLY AT
CC CELL-CELL CONTACT SITES AND IS CONTINUOUSLY FOUND IN THE RETINAL
CC AXON TERMINAL ARBOR LAYERS OF THE ADULT TECTUM.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like V-type domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----

```

```

DR EMBL, L25056; AAC38015.2; -.
DR PIR, I50478; I50478.
DR HSSP, Q13740; IKTIC.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00406; Ig; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 1.
DR Signal; Developmental protein; Cell adhesion; Immunoglobulin domain;
KM Repeat; Glycoprotein; Transmembrane.
FT SIGNAL 1 22
FT CHAIN 23 555
FT DOMAIN 23 555
FT TRANSMEM 500 520
FT DOMAIN 521 555
FT DOMAIN 31 127
FT DOMAIN 131 229
FT DOMAIN 239 323
FT DOMAIN 319 397
FT DOMAIN 406 484
FT DISULFID 38 110
FT DISULFID 154 217
FT DISULFID 263 306
FT CARBOHYD 92 92
FT CARBOHYD 171 171
FT CARBOHYD 350 350
FT CARBOHYD 441 441
FT CARBOHYD 465 465
SQ SEQUENCE 555 AA; 60371 MW; 5AAB8014F08BFF68 CRC64;

Query Match 5.0%; Score 116; DB 1; Length 555;
Beet Local Similarity 19.0%; Pred. No. 1.3;
Matches 75; Conservative 69; Mismatches 129; Indels 122; Gaps 20;

OY 29 VLKRGDPIVELTCTASOKS---IQPHKNSNOIKILNOSGFLTKGSK-----LNDK 79
DB 26 VIGLGETIVPCNDGTKKPDGLITTKM--YVDDSPGDLVKAQKQKQKATVSATDG 82
OY 80 ADSRSLWDQGNFPIIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLQGSGL 139
DB 83 YKSRVST--ANSSILIRAGSLADRVFTC-----MVSFTN-----LEESV 123
OY 140 TLTLESPPGSSPSVOCSPRGKNIQGTLSVQLELQDSGTWCTVLQNKVFEKIDI 199
DB 124 EVKVKKRP-SARVIR--NNAKELNGKLTQGECEVENA----- 159
OY 200 VVLAQKASSTVYKKEGEV-----EFSFPLAFVTEKLTGSGELMWQAEKASSKSWITF 254
DB 160 ----NPPADLIWKKNQTLVDGKTIITSTITDKITG-----ISSTSSRLQY 204
OY 255 DLKNEKVSRYKVTQDPKQLQMGKCLPLHLTLPOALQVYAGSGNLTALAEKTKGLH---OE 311
DB 205 TAKKEDVE-SQFTCTAKHYMG--PDQVSEPSFP-----IHYPTK 242
OY 312 VNLVVMRAITOLQK---NLTCVWGPTSPSKMLSLKLENKAKYSKREKPVWL---NP 363
DB 243 VSLQVVSQSPIREGKVDVTLKQCADGNPP---TSTFNFNKKGKGVTVTKDQVYTLTGVTRA 299
OY 364 EAGMWQCLSDSGQVLLS-----NIKVLPT 389
DB 300 DSGIYKCSLLDND--VMESTQFVTVSFLDVSILPT 332

```

```

DE Neural cell adhesion molecule 1, 140 kDa isoform precursor (N-CAM 140)
GN (NCAM-140).
DE NCAM1 OR NCAM.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid:9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain cortex;
RX MEDLINE=89378239; PubMed=2776887;
RA Lipkin V.M., Khramtsov N.V., Andreeva S.G., Moshayakov M.V.,
RA Petukhova G.V., Rakitina T.V., Reshchenko E.A., Ishchenko K.A.,
RA Mitroeva S.F., Chernova M.N., Dzenytsyna S.M.;
RT "Calmodulin-independent bovine brain adenylyl cyclase. Amino acid
RL FEBS Lett. 284:69-73(1989).
RN [2]
RP SEQUENCE OF 20-36.
RX MEDLINE=86140120; PubMed=3512556;
RA Rougon G., Marshak D.R.;
RT "Structural and immunological characterization of the amino-terminal
RL J. Biol. Chem. 261:3396-3401(1986).
RN [3]
RP IDENTIFICATION AS N-CAM.
RX MEDLINE=92111748; PubMed=1765159;
RA Premont R.T.;
RT "A bovine brain cDNA purported to encode calmodulin-insensitive
RL adenylyl cyclase has extensive identity with neural cell adhesion
RT molecules (N-CAMs).";
RL FEBS Lett. 295:230-231(1991).
CC -1- FUNCTION: This protein is a cell adhesion molecule involved in
CC neuron-neuron adhesion, neurite fasciculation, outgrowth of
CC neurites, etc.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=1;
CC Comment=A number of isoforms are produced;
CC Name=1;
CC IsoId=P31836-1; Sequence=Displayed;
CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
CC -1- CAUTION: Was originally (Ref.1) thought to be a calmodulin-
CC independent adenylyl cyclase.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL, X16451; CA34470.1; -.
DR PIR, A32976; IJBONC.
DR HSSP, P40189; IBOU.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00408; IGc2; 5.
DR PROSITE; PS50835; IG_LIKE; 5.
KM Cell adhesion; Glycoprotein; Transmembrane; Repeat;
KM Immunoglobulin domain; Alternative splicing; Signal.
FT SIGNAL 1 19
FT CHAIN 20 853
FT DOMAIN 20 719
EXTRACELLULAR (POTENTIAL).

```

```

FT TRANSMEM 720 737 POTENTIAL.
FT DOMAIN 738 853 CYTOPLASMIC (POTENTIAL).
FT 20 111 IG-LIKE C2-TYPE 1.
FT DOMAIN 116 305 IG-LIKE C2-TYPE 2.
FT 212 300 IG-LIKE C2-TYPE 3.
FT DOMAIN 307 412 IG-LIKE C2-TYPE 4.
FT 415 500 IG-LIKE C2-TYPE 5.
FT DOMAIN 527 604 FIBRONECTIN TYPE-III 1.
FT 633 700 FIBRONECTIN TYPE-III 2.
FT DOMAIN 152 156 HEPARIN-BINDING (POTENTIAL).
FT 161 165 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 41 96 BY SIMILARITY.
FT DISULFID 139 189 BY SIMILARITY.
FT DISULFID 235 286 BY SIMILARITY.
FT DISULFID 328 394 BY SIMILARITY.
FT 435 488 BY SIMILARITY.
FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 346 346 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 853 AA; 93893 MW; 512FD9231A7A368 CRC64;

```

```

Query Match 5.0%; Score 116; DB 1; Length 853;
Best Local Similarity 20.5%; Pred.No. 2.2;
Matches 87; Conservative 52; Mismatches 159; Indels 126; Gaps 18;

```

```

OY 32 KKQDVELTCTASQKKSIOFMKNSNOIKILNGSGSLTKPSKLNDRASRLMDQGN 91
DB 130 REGDAVAVICVSVSLPPTIIMKHGRDVIITKKDVRFTV-----LTN 171
OY 92 PPLIINKIKIDSDTYICE-----VEDQKEVOLV-----FGLTANSDTHL 133
DB 172 NYLQIRGKIKKTDEGYRCEGRILARGEINFKDIQYIVAVPPTVQAROSIVATAN----- 226
OY 134 LOGGSLTTLTSPGSSPSVOCSPRGKNIQG-----GKTLVSQLELDQSGTW 182
DB 227 -LGOSVTLVCAEGFPEPTVGM-TWQEQIENEDEKTLFSDDSSELTIRKVDKNDEAY 284
OY 183 TCTVLQNGKVKYFKDIYVLAFOKASSIVYKKEGQVFSPLATYKLTGSGE----- 237
DB 285 VCIENKAGQEDASIHLYKFKPK--ITVYENQTMALBEOVLTTC--ASDDPIPSI 338
OY 238 LMMQERASSSK--SWITFDLKNKEVSKRVTOPKIQMGKKPL-HLTLPQALPOYAG 293
DB 339 TWRSTRISSSEBRKSWTRPE-----KQETLDGHVAVSHARVSLTIKSI--QYTD 388
OY 294 SGNLTTLAETKGLHQEVNLVMEKATQL-----QKNLTCEVWGPTSPKMLLS 341
DB 389 AGEVYCTASNTIGDQSGMYLEVQYAPKLQCPVAVYTWEGNQVNITCEVFAYPS----- 442
OY 342 LKLENKEKVKREKRPVAVNLPEAGMWQCLSDSQVLEBNIKVLPFWSPVPRASAL 401
DB 443 -----ATIS-----WFRD-----GQLLPSSNYSNIKIYNT--PSASYL 473
OY 402 PAPP 405
DB 474 EVTP 477

```

```

RESULT 54
FPRP RAT STANDARD; PRT; 879 AA.
AC Q62786;

```

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Prostaglandin F2 receptor negative regulator precursor (Prostaglandin
DE F2-alpha receptor regulatory protein) (Prostaglandin F2-alpha receptor
DE associated protein).
GN PTPFRN OR FPRP.
OS Rattus norvegicus (Rat).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Ovary;
RX MEDLINE=97109715; PubMed=8951995;
RA Orlicky D.J., Nordeen S.K.;
RT "Cloning, sequencing and proposed structure for a prostaglandin F2
RT alpha receptor regulatory protein."; Acids 55:261-268 (1996).
RL Prostaglandins leukot. Essent. Fatty Acids 55:261-268 (1996).
RN [2]
RN CHARACTERIZATION.
RP STRAIN=Sprague-Dawley; TISSUE=Ovary;
RX MEDLINE=96397016; PubMed=8804121;
RA Orlicky D.J.;
RT "Negative regulatory activity of a prostaglandin F2 alpha receptor
RT associated protein (FPRP).";
RL Prostaglandins leukot. Essent. Fatty Acids 54:247-259 (1996).
CC -1- FUNCTION: Inhibits the binding of prostaglandin F2-alpha (PGF2-
CC alpha) to its specific FP receptor, by decreasing the receptor
CC number rather than the affinity constant. Functional coupling with
CC the prostaglandin F2-alpha receptor seems to occur.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
CC reticulum and trans-Golgi network.
CC -1- TISSUE SPECIFICITY: REPRODUCTIVE TISSUES, LUNG AND HEART.
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U26595; AAC18426.1; -
DR GO; GO:0005515; F:Protein binding; IPL.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 6.
DR SMART; SM00406; IGV; 1.
DR POSITIVE; PSS0835; IG LIKE; 5.
KW Signal; Immunoglobulin domain; Transmembrane; Glycoprotein; Repeat;
KW Polymorphism.
FT SIGNAL 1 21
FT CHAIN 22 879
FT FT
FT DOMAIN 22 832
FT TRANSMEM 833 853
FT 854 879
FT DOMAIN 22 129
FT 149 268
FT DOMAIN 276 389
FT 406 527
FT DOMAIN 544 662
FT 688 813
FT DOMAIN 424 427
FT 89 91
FT SITE 703 705
FT DISULFID 43 119
FT DISULFID 169 247
FT DISULFID 299 373
FT DISULFID 429 515
FT DISULFID 571 793
FT CARBOHYD 44 44
FT CARBOHYD 300 300
FT CARBOHYD 383 383
FT CARBOHYD 413 413
FT CARBOHYD 525 525
FT CARBOHYD 600 600

```

FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 691 691 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 47 47 D -> G.
FT VARIANT 136 136 M -> V.
FT VARIANT 782 782 S -> G.
FT VARIANT 844 844 G -> R.
SQ SEQUENCE 879 AA; 98730 MW; DE7012D3B346C0F7 CRC64;

Query Match 5.0%; Score 115.5; DB 1; Length 879;
Best Local Similarity 18.6%; Pred. No. 2.5;
Matches 87; Conservative 78; Mismatches 175; Indels 129; Gaps 18;

14 LQALPAPATQGNKVLGKGDVTELTCTASQ-----KKSQPFMKNSNQKILG----- 63
10LAV-----PRVSVTEGKDLDSCTITDRDVAPEVTWTFKKTPTSLASHML 331
64 ---NGSFLLTKGPSKLNRADRSRLMDGNFLLIKKLEDSPTYICEV-----ED 113
332 ARLDRLSLVHSSPHVALSHVDTR-----SYHLVRDVSKEKNGVYLCVLAMPGNHR 384
QY 114 QKEEVQLVFLGLTANSSTHLLGOSLTLTLESPPSS---PSVQCRSPRKNIGGKTLIS 170
DB 385 SMHKAEMASAPSGVSVTLBEPEOVYLNASKVPFSPDPTELOCRVIDTKKVDAGVRLT 444
QY 171 V-----SOLEQDSGTWCTVLONQKKVEFKIDIVLAFOKA--SIIYK 213
DB 445 VSMYRMRNRNDVVASSELLAMVDGWTIRYERSK-----QRQDSEPFIS 491
QY 214 KEGQVEFSFPLAFVTEKLTGSGELMWQERASSSSKSWI--TPDLKNKEVSVKRVTDPTL 272
DB 492 KENTDT-FSFRIGRTTEEDRSGSYCVASAMTRQNRSSWKSADVPSKPNIFWASEDSVL 550
QY 273 QMGKPLPLHLTPALPOYASGSLTLLEATGKTLHGVNLVMPATQLQKNLTCEWVG 332
DB 551 VKKAP-----KPFPAAGNTEPMCKVSSKNIKSPRYSVLITAEKPVGDL-----S 598
QY 333 PTPSKLMLSL-----KLEN-----KEAKYSKREKPPVNLNPEAGMOCILS 373
DB 599 PNETYITISLDDSVYKLEKMTDASRVDCVLEKQVDEFFRYKRWQTOVSQDGLYRCWVT 658
QY 374 DSGQVLLBSNIVKLPSTFVPRASALPAPPTGSALEPPQPTASALPP 422
DB 659 -----AMS-----PIGSL-WREATSLSN 678

RESULT 55
TIE1 BOVIN
ID TIE1 BOVIN STANDARD; PRT; 1136 AA.
AC 006805;
DT 01-FEB-1995 (rel. 31, Created)
DT 01-FEB-1995 (rel. 31, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DE Tyrosine-protein kinase receptor Tie-1 precursor (EC 2.7.1.112).
GN TIE OR TIE1 OR TIE-1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Kuminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Endothelial cells;
RX MEDLINE=9402374; PubMed=8415706;
RA Sato T.N., Qin Y., Kozak C.A., Andus K.L.;
RT "Tie-1 and tie-2 define another class of putative receptor tyrosine
RT kinase genes expressed in early embryonic vascular system.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:9355-9358(1993).
CC -1- FUNCTION: Probable protein tyrosine-kinase transmembrane receptor.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Specifically expressed in developing vascular
CC endothelial cells.

CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Tie
CC subfamily.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 3 fibronectin type III domains.
CC -1- SIMILARITY: Contains 3 EGF-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

EMBL; X71423; CA50554.1; -.
PIR; S57845; S57845.
DR HSSP; P11362; IEGK.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR002049; Laminin EGF.
DR InterPro; IPR000719; Prot. kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00047; Ig; 2.
DR Pfam; PF00041; Fn3; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00409; IG; 1.
DR SMART; SM00219; TYKIC; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Receptor, Tyrosine-protein kinase; Transferase; Signal; ATP-binding;
KW Repeat; EGF-like domain; Transmembrane; Immunoglobulin domain;
KW Glycoprotein; Phosphorylation; Multigene family.
FT SIGNAL 1 23
FT CHAIN 24 1136
FT DOMAIN 24 757
FT TRANSMEM 758 782
FT DOMAIN 783 1136
FT DOMAIN 43 106
FT DOMAIN 212 254
FT DOMAIN 256 301
FT DOMAIN 303 343
FT DOMAIN 370 424
FT DOMAIN 444 538
FT DOMAIN 541 637
FT DOMAIN 641 742
FT DOMAIN 837 1116
FT NP_BIND 843 851
FT BINDING 868 868
FT ACT_SITE 977 977
FT MOD_RES 1005 1005
FT CARBOHYD 84 84
FT CARBOHYD 159 159
FT CARBOHYD 501 501
FT CARBOHYD 594 594
FT CARBOHYD 707 707
SQ SEQUENCE 1136 AA; 124953 MW; EEF85804A041BB12 CRC64;

Query Match 5.0%; Score 115.5; DB 1; Length 1136;

| | | | | | |
|----|--|---|---|--|---|
| CC | | | | | NEURONS AND GLIA IN THE DEVELOPING NERVOUS SYSTEM AND THE SHORT |
| CC | | | | | ISOPROB TO OTHER NONNEUROBL TISSUES. |
| CC | -1- | STIMILARITY: | Contains 6 Immunoglobulin-like C2-type domains. | | |
| CC | -1- | SMILARITY: | Contains 5 fibronectin type III domains. | | |
| CC | | | | | ----- |
| CC | This | SWISS-PROT entry | is copyright. It is produced through a collaboration | | |
| CC | between | the Swiss Institute of Bioinformatics | and the EMBL outstation - | | |
| CC | the European Bioinformatics Institute. | There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | | |
| CC | modified and this statement is not removed. | Usage by and for commercial | | | |
| CC | entities requires a license agreement (see http://www.isb-sib.ch/announce/ | | | | |
| CC | or send an email to license@isb-sib.ch). | | | | |
| CC | | | | | ----- |
| DR | EMBL; | M28231; AAA28728.2; - | | | |
| DR | EMBL; | AF050085; AAC28613.2; - | | | |
| DR | EMBL; | AF050084; AAC28613.2; JOINED. | | | |
| DR | EMBL; | AF050085; AAC28614.2; - | | | |
| DR | EMBL; | AF050084; AAC28614.2; JOINED. | | | |
| DR | EMBL; | AEO03444; AAP46387.1; - | | | |
| DR | EMBL; | AY058284; AAL13513.1; - | | | |
| DR | EMBL; | X76243; CAAS3822.1; - | | | |
| DR | PIR; | A32579; A32579. | | | |
| DR | PDB; | 1CFB; 30-NOV-94. | | | |
| DR | FlyBase; | FBgn0002968; Nrg. | | | |
| DR | GO; | GO:0005886; C:plasma membrane; IMP. | | | |
| DR | GO; | GO:0005194; P:cell adhesion molecule activity; IMP. | | | |
| DR | GO; | GO:0007560; P:marginal disc morphogenesis; IMP. | | | |
| DR | InterPro; | IPR008957; FN_III-like. | | | |
| DR | InterPro; | IPR003961; FN_III. | | | |
| DR | InterPro; | IPR007110; Ig-I like. | | | |
| DR | Pfam; | PFO0041; fn3; 5. | | | |
| DR | Pfam; | SFO0047; Ig; 6. | | | |
| DR | SMART; | SMO0060; FN3; 5. | | | |
| DR | SMART; | SMO0408; IGC2; 4. | | | |
| DR | PROSITE; | PS50835; IG_LIKE; 6. | | | |
| KM | Cell adhesion; Glycoprotein; Transmembrane; Repeat; 3D-structure; | | | | |
| KM | Immunoglobulin domain; Signal; Developmental protein; | | | | |
| KM | Alternative splicing. | | | | |
| FT | SIGNAL | 1..23 | | | |
| FT | CHAIN | 24..1302 | NEUROGLIAN. | | |
| FT | DOMAIN | 24..1338 | EXTRACELLULAR (POTENTIAL). | | |
| FT | TRANSMEM | 1139..1154 | POTENTIAL. | | |
| FT | DOMAIN | 1155..1302 | CYTOPLASMIC (POTENTIAL). | | |
| FT | DOMAIN | 29..133 | IG-LIKE C2-TYPE 1. | | |
| FT | DOMAIN | 134..225 | IG-LIKE C2-TYPE 2. | | |
| FT | DOMAIN | 245..330 | IG-LIKE C2-TYPE 3. | | |
| FT | DOMAIN | 339..426 | IG-LIKE C2-TYPE 4. | | |
| FT | DOMAIN | 432..524 | IG-LIKE C2-TYPE 5. | | |
| FT | DOMAIN | 521..610 | IG-LIKE C2-TYPE 6. | | |
| FT | DOMAIN | 629..690 | FIBONECTIN TYPE-III 1. | | |
| FT | DOMAIN | 729..792 | FIBONECTIN TYPE-III 2. | | |
| FT | DOMAIN | 832..896 | FIBONECTIN TYPE-III 3. | | |
| FT | DOMAIN | 932..997 | FIBONECTIN TYPE-III 4. | | |
| FT | DOMAIN | 1024..1098 | FIBONECTIN TYPE-III 5. | | |
| FT | DISULFID | 59..111 | POTENTIAL. | | |
| FT | DISULFID | 625..706 | | | |
| FT | CARBOHYD | 182..182 | N-LINKED (GLCNAC . .) (POTENTIAL). | | |
| FT | CARBOHYD | 198..198 | N-LINKED (GLCNAC . .) (POTENTIAL). | | |
| FT | CARBOHYD | 411..411 | N-LINKED (GLCNAC . .) (POTENTIAL). | | |
| FT | CARBOHYD | 448..448 | N-LINKED (GLCNAC . .) (POTENTIAL). | | |
| FT | CARBOHYD | 652..652 | N-LINKED (GLCNAC . .) (POTENTIAL). | | |
| FT | CARBOHYD | 683..683 | N-LINKED (GLCNAC . .) (POTENTIAL). | | |
| FT | CARBOHYD | 821..821 | N-LINKED (GLCNAC . .) (POTENTIAL). | | |
| FT | CARBOHYD | 1125..1125 | N-LINKED (GLCNAC . .) (POTENTIAL). | | |
| FT | VARSPLIC | 1224..1239 | OFTEDGSFIGGVPRK -> MNEDGSFIGGGRKL (in isoform Short). | | |
| FT | /FTid=VSP_002601. | | | | |
| FT | Missing (In Isoform Short). | | | | |
| FT | /FTid=VSP_002602. | | | | |
| FT | NR -> KP (IN REF. 2). | | | | |
| FT | CONFLICT | 85..86 | | | |

| | | | | | |
|----|----------|------|------|--|---------------------|
| FT | CONFLICT | 1282 | 1282 | | MISSED (IN REF. 6). |
| FT | STRAND | 619 | 625 | | |
| FT | STRAND | 629 | 635 | | |
| FT | TURN | 640 | 641 | | |
| FT | STRAND | 646 | 653 | | |
| FT | TURN | 657 | 658 | | |
| FT | STRAND | 661 | 668 | | |
| FT | TURN | 669 | 670 | | |
| FT | STRAND | 673 | 677 | | |
| FT | STRAND | 682 | 692 | | |
| FT | TURN | 693 | 694 | | |
| FT | STRAND | 695 | 696 | | |
| FT | STRAND | 706 | 708 | | |
| FT | STRAND | 721 | 723 | | |
| FT | TURN | 727 | 728 | | |

Query Match 5.0%; Score 115.5; DB 1; length 1302;
Best Local Similarity 19.7%; Pred. No. 4.2; Indels 83; Gaps 16
Matches 67; Conservative 61; Mismatches 129;

| | | |
|----|-----|---|
| Dy | 25 | GNNKVLGGKKGVLTCTTASQK--SIOFHKNSNQIKLGNQS-FLTKGPSKLNDRA-80 |
| Dy | 225 | GNNKVL-----DYKQMGVSAQNKHPPRVQYVSRRGSLRLRGKRMELFCIYGSTPLPQTW-280 |
| Dy | 81 | --DSRRSLWD----QGNF-PLIIKNLIKEDSDTYICEVEDQKEVOLIVFGLTANSOHT-132 |
| Dy | 281 | SKDGRIGWSRITGTGHGKSLVRKQTNFDAGTTCIDVSNCGNAOSFSILTNVNSVPY-340 |
| Dy | 133 | LLOGSLLTLTESPRGSSPSVQCS---PRGNKIQSGK-----TLSTV-171 |
| Dy | 341 | FTKEPEITAAED--EEVFECRAGVPEPKISWIHNKGPTEOSTPNPRRTVTNNTIKI-397 |
| Dy | 172 | SQLELDSDGTWTCTVLON----QKKVEFKIDIIVLAFOKASSIVYKKEGEQVESFPILA-227 |
| Dy | 398 | INLVKGDPTGNGVCNANTSLGYVKDVLYLNQAEPPTISEAPPAVSTVDGRNV-----449 |
| Dy | 228 | TVE-KLTGSGELMWQAEPASSSKSWIT--FDLK-NKEVSVKRYTDPEPLQMGKKLPHL-282 |
| Dy | 450 | TIKCRVNSPKPLVWLIRAS--NWLTGGRVYNQANGDLIEDVLT-----491 |
| Dy | 283 | TLPOALPOYAGSGNLTLALEKTGKLGHOEVNLVYWKATQL-322 |
| Dy | 492 | -----FSDAGKYTCYAQNKFGEIQADGSLVYKHEHTRI-523 |

RESULT 57
FCEA_HUMAN STANDARD; PRT; 257 AA.
AC FCEA_HUMAN P12319;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE High affinity immunoglobulin epsilon receptor alpha-subunit precursor (FcER1) (IGE FC receptor, alpha-subunit) (Fc-epsilon RI-alpha).
DE DE FCBRIA OR FCBIA.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RP MEDLINE=88233953; PubMed=2967464;
RA Kochan J., Pettine L.F., Hakimi J., Kishi K., Kinet J.-P.;
RT "Isolation of the gene coding for the alpha subunit of the human high
RT affinity IGE receptor.";
RL Nucleic Acids Res. 16:3584-3584(1988).
[2]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Mast cells;
RX MEDLINE=88158102; PubMed=2964640;
RA Shimizu A., Teplov I., Bentley P.N., Berenstein E.H., Siraganian R.P.,
RA Leder P.;
RT "Human and rat mast cell high-affinity immunoglobulin E receptors:

| | |
|----|--|
| RT | characterization of putative alpha-chain gene products."; |
| RL | Proc. Natl. Acad. Sci. U.S.A. 85:1907-1911(1988). |
| RP | 3D-STRUCTURE MODELING OF 26-197. |
| RX | MEDLINE=93113350; PubMed=1472946; |
| RA | Padian E.A., Heim B.A.; |
| RT | "A modeling study of the alpha-subunit of human high-affinity receptor for immunoglobulin-E."; |
| RT | Receptor 2:129-144(1992). |
| CC | -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS EPSILON. HIGH AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE) RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES. |
| CC | -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO DISULFIDE LINKED GAMMA CHAINS. |
| CC | -1- SUBCELLULAR LOCATION: Type 1 membrane protein. |
| CC | -1- SIMILARITY: Contains 2 immunoglobulin-like domains. |
| CC | ----- |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). |
| CC | ----- |
| DR | EMBL; X06948; CAA30025.1; - |
| DR | EMBL; J03505; AAA36204.1; - |
| DR | EMBL; A21606; CAA01564.1; - |
| DR | PIR; S00682; S00682. |
| DR | PDB; 1ALS; 27-FEB-95. |
| DR | PDB; 1ALT; 27-FEB-95. |
| DR | PDB; 1F2Q; 08-JUN-00. |
| DR | PDB; 1J86; 29-AUG-01. |
| DR | PDB; 1J87; 29-AUG-01. |
| DR | PDB; 1J88; 29-AUG-01. |
| DR | PDB; 1J89; 05-SEP-01. |
| DR | Genew; HGNC:3609; FCER1A. |
| DR | MIM; 147140; - |
| DR | GO; GO:0005887; C:integral to plasma membrane; TAS. |
| DR | InterPro; IPR007110; IG-1-like. |
| DR | InterPro; IPR003598; IG_1c2. |
| DR | Pfam; PF00047; Ig; 2. |
| DR | SMART; SM00408; IGC2; 1. |
| DR | PROSITE; PS50835; IG_LIKE; 2. |
| KW | IgF-binding protein; Receptor; Transmembrane; Glycoprotein; Signal; Immunoglobulin domain; Repeat; 3d-structure. |
| FT | SIGNAL |
| FT | CHAIN |
| FT | 26 257 |
| FT | 26 205 |
| FT | DOMAIN |
| FT | TRANSSEM |
| FT | 206 224 |
| FT | DOMAIN |
| FT | 225 257 |
| FT | 30 110 |
| FT | DOMAIN |
| FT | 111 193 |
| FT | DISULFID |
| FT | 51 93 |
| FT | DISULFID |
| FT | 132 176 |
| FT | CARBOHYD |
| FT | 46 46 |
| FT | CARBOHYD |
| FT | 67 67 |
| FT | CARBOHYD |
| FT | 75 75 |
| FT | CARBOHYD |
| FT | 99 99 |
| FT | CARBOHYD |
| FT | 160 160 |
| FT | CARBOHYD |
| FT | 165 165 |
| FT | CARBOHYD |
| FT | 191 191 |
| FT | STRAND |
| FT | 31 35 |
| FT | HELIX |
| FT | 39 42 |
| FT | TURN |
| FT | 43 43 |
| FT | STRAND |
| FT | 48 54 |
| FT | STRAND |
| FT | 61 66 |
| FT | TURN |
| FT | 67 68 |
| FT | STRAND |
| FT | 69 71 |

[illegible]

```

RN [2]
RP REVISIONS.
RA Kochan J.;
RL Submitted (MAR-1988) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Mast cells;
RX MEDLINE=88158102; PubMed=2964640;
RA Shimizu A., Tepler I., Bentley P.N., Berenstein E.H., Siraganian R.P.,
  Leder P.;
RT "Human and rat mast cell high-affinity immunoglobulin E receptors:
  characterization of putative alpha-chain gene products.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:1907-1911(1988).
RN [4]
RP SEQUENCE OF 21-245 FROM N.A.
RX MEDLINE=88289772; PubMed=2969594;
RA Liu F.-T., Abdrant K., Robertson M.W.;
RT "CDNA heterogeneity suggests structural variants related to the high-
  affinity IgE receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:5639-5643(1988).
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
  AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
  RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IgE LEADS TO CELL
  ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
  RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
  ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
  DISULFIDE LINKED GAMMA CHAINS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M17153; AAA42045.1; -
DR EMBL, J03606; AAA41582.1; -
DR EMBL, M21622; AAA41146.1; -
DR PIR, C31327; A30154.
DR HSSP; P12319; 1ALS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR Pfam; PF00047; Ig_2.
DR SMART; SM00409; Ig; 2.
DR PROSITE; PS00835; Ig_Like; 1.
KW IgE-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
  Immunoglobulin domain; Repeat.
KW SIGNAL.
FT CHAIN 1 23
FT 24 245 HIGH AFFINITY IMMUNOGLOBULIN EPSILON
  RECEPTOR ALPHA-SUBUNIT.
FT DOMAIN 24 204 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 205 223 POTENTIAL.
FT DOMAIN 224 245 CYTOPLASMIC (POTENTIAL).
FT 28 103 IG-LIKE 1.
FT DOMAIN 113 181 IG-LIKE 2.
FT DISULFID 49 91 BY SIMILARITY.
FT 130 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 52 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 245 AA; 27793 MW; A0867D363B72197 CRC64;
Query Match 4.9%; Score 113.5; DB 1; Length 245;
Best Local Similarity 23.5%; Pred. No. 0.66;
Matches 63; Conservative 41; Mismatches 85; Indels 79; Gaps 14;

```

```

QY 11 LVLQLALLPAATGKNVVLGK-----KEDVELTC---TASQKSIQPHKNSNQIKI 61
DB 13 LVLLSLGMLTATKQSVSLDPPWIRITLTGKTLICGNNSQMSNRK--WINDSL-- 68
QY 62 LGNGSPLTKGPSKLNDRASRRSLWDQGNPFLIKIKLIEDSDPYICEVED--QKEVQ 119
DB 69 -----SNVKSASH-----YIVSATIIDSGKYLQKGFYKSKFVY 103
QY 120 LLYEGLTANSTHLGQSLTLTLESPPGSSPSVQCSPRG-----KTIQ 164
DB 104 LNMV-----QEWLLQSSADVLDN--GSFDIRCSWKWKXKVKYIKDIAFKKSY 154
QY 165 GKTLSVQSLQLELOSGTWTCTVLONOKKVEFKIDIVLAFOKASSIYKKEGEVFEFP 224
DB 155 DSNNISIKATPNDSGSHCTGYLN--KVECKSDKFSIAVVKDYTIETR---WLQLIFP 208
QY 225 ----LAFTVEKLTGSGELMWQAEKASSS 248
DB 209 SLAVILFAVD--TG---LWFSTHQFES 231

RESULT 59
LAMP CHICK STANDARD; PRT; 338 AA.
AC 098919;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Limbic system-associated membrane protein precursor (E19S) (CHLAMP,
  G19-isoform).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97157768; PubMed=9004047;
RA Wilson D.J.A., Kim D.S., Clarke G.A., Marshall-Clarke S., Moss D.J.;
RT "A family of glycoproteins (GP55), which inhibit neurite outgrowth,
  are members of the Ig superfamily and are related to OBGM,
  neurotrophin, LAMP and CEP-1.";
RL J. Cell Sci. 109:3129-3138(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97358596; PubMed=9215692;
RA Brummendorf T., Spaltmann F., Treubert U.;
RT "Cloning and characterization of a neural cell recognition molecule
  on axons of the retinorectal system and spinal cord.";
RL Eur. J. Neurosci. 9:1105-1116(1997).
CC -1- FUNCTION: Mediates selective neuronal growth and axon targeting.
  Probably serves as a recognition molecule for the formation of
  limbic connections (By similarity)
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
  similarity).
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. IGLON
  family.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL, Y08171; CAA69357.1; -
DR EMBL, Z94720; CAB08115.1; -
DR InterPro; IPR007110; Ig-like.

```

| | |
|----|--|
| DR | InterPro: IPR003598; Ig_c2. |
| DR | Pfam: PF00047; ig_3. |
| DR | SMART: SM00408; IGc2_2. |
| DR | PROSITE: PS50833; IG_LIKE_3. |
| KW | Immunoglobulin domain; Cell adhesion; Glycoprotein; Lipoprotein; |
| KW | Membrane; GPI-anchor; Repeat; Signal. |
| FT | SIGNAL 1 28 POTENTIAL. |
| FT | CHAIN 29 315 LIMBIC SYSTEM-ASSOCIATED MEMBRANE PROTEIN. |
| FT | PROPEP 316 338 REMOVED IN MATURE FORM (POTENTIAL). |
| FT | DOMAIN 29 122 IG-LIKE C2-TYPE 1. |
| FT | DOMAIN 132 214 IG-LIKE C2-TYPE 2. |
| FT | DOMAIN 219 306 IG-LIKE C2-TYPE 3. |
| FT | DISULFID 53 111 POTENTIAL. |
| FT | DISULFID 153 197 POTENTIAL. |
| FT | DISULFID 239 290 POTENTIAL. |
| FT | LIPID 315 GPI-anchor amidated asparagine (Potential). |
| FT | CARBOHYD 40 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD 66 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD 136 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD 148 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD 279 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD 287 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD 300 N-LINKED (GLCNAC. .) (POTENTIAL). |
| FT | CARBOHYD 315 N-LINKED (GLCNAC. .) (POTENTIAL). |
| QO | SEQUENCE 338 AA; 37394 MW; 8FAA460AD9842EB4 CRC64; |

| | | | | |
|-----------------------|--------|--------------|-------|-----------------|
| Query Match | 4.9%; | Score 113.5; | DB 1; | Length 338; |
| Best Local Similarity | 21.0%; | Pred. No. 1; | | |
| Matches | 83; | Conservative | 64; | Mismatches 144; |
| | | | | Indels 105; |
| | | | | Gaps 20; |

```

Oy      10 LLLVLQDALLPAA-----TOGKKVVLGGKKGPVETLTCAQOKKSIOFMKKNIOIKI 61
Db      14 LVLLRLCLLPGLPVRVSDFTRGDINTVIRGSDTALLRCVEEDRSS-KAAMLRSGITIF 72

Oy      62 LGNQGSFLTQKPSKLNDRAD-SRRLSDMGNPILLIKLKIEDSDPIYCEVEDO---KE 116
Db      73 AG-----EDKMSLDPRVELEKRSPL-YSALKQVDVYDGSSTYCSVQTHHPKTS 122

Oy      117 EVOLLVFG---LTANSDTHLLQGSULTLTSPGSSPSVQCR-SPRKNIOGKGT-L 169
Db      123 QVVLIVQVPKPSKNISSTDITVEGSSVTLVCMANGRPPIVTHRHLLPTQGEPEGEBEYL 182

Oy      170 SVSQGLELDQSGTWTCTVLQNKQKVERKIDIVLARQKASIVYKKKEBOYE-FSPFLAF 227
Db      183 EIIGITREBSGKYEC-----KAANEVASADVAQVAVTVNVPPTI 221

Oy      228 TVEKLTGSGELMWQARASSKSWITFDLKNKEVSV-----KRVTDPRKLQMGKPLRH 281
Db      222 TESK-----SNBAATGRQA-----LLREBAVAFTPPPEWYRDDTRINSANGLEIK 267

Oy      282 LTLPLPALPOYASGNITLALAEKTKGLHOEVLNVLMRATOLQOKULTCEWGSPTSPKML- 340
Db      268 ST-----GSGSLIMVANVATVEEHYG-NYTCVAANKL-----GVTVASLYLY 306

Oy      341 ---SLKLENKEAVKSKREKPVVNLVLPBAGMGCILS 373
Db      307 RPGTGRVDNGSVSLA---VPLMLL---PASTLLCILIS 336

```

| RESULT 60 | CAML_FUGRU | STANDARD: | PRT: 1277 AA. |
|-----------|---|-----------|---------------|
| ID | CAML_FUGRU | | |
| AC | Q98902; | | |
| DT | 10-OCT-2003 (rel. 42, Created) | | |
| DT | 10-OCT-2003 (rel. 42, Last sequence update) | | |
| DT | 15-MAR-2004 (rel. 43, Last annotation update) | | |
| DE | Neural cell adhesion molecule L1 precursor (N-CAM L1) (L1-CAM). | | |
| GN | L1CAM. | | |
| OC | Pingu ruidripes (Japanese pufferfish) (Takifugu rubripes). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei | | |

[illegible]

```

FT DISULFID 554 607 BY SIMILARITY.
FT CARBOHYD 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 503 503 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 520 520 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 794 794 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 839 839 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1035 1035 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1046 1046 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1068 1068 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1083 1083 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1108 1108 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT VARSPLIC 42 47 VYISDL -> F (in isoform 2).
FT VARSPLIC 1190 1193 /FTid=VSP_050474.
SQ SEQUENCE 1277 AA; 141954 MW; 2848B49BA9A42C27 CRC64;
Query Match 4.9%; Score 113.5; DB 1; Length 1277;
Best Local Similarity 21.7%; Pred. No. 5.6;
Matches 80; Conservative 56; Mismatches 152; Indels 81; Gaps 16;
QY 34 GDTVELTCTASQKSIQFHKNSNOIKILNGSFLTKGPSKLNDRADSRSLMDQGNFP 93
DB 363 GETVLDLCOADGIPSPITW-----TVNGVPLSATSLPRRSITLSSGS-- 405
QY 94 LIKKLKIEDSTYICEVEDQEE-----VOLVPLGLTANSDTH-LLOGSLTL-- 141
DB 406 LILKQVIFEDDTAIYCCQASNKKGITLANTNVYIELPQILLENENTYTFVGGOKALLEC 465
QY 142 -TLSP-----PGSSPSVQCRSPRKNIQGGKTSVLSQELDQSGTWCTYLQKQKYE 194
DB 466 ETRGSPKRVNTWESSISILLADPRVNLITNG-GEIANSVHDDSGITCLVQGSNISVN 524
QY 195 FKIDI-----VLAFOKASSIVYKKEGVEFSFPPLAFVEXKLTGSGELMWAQER--ASS 247
DB 525 AVEVLNRTVILSPQALRL--QPKTAIFTC-LYVTDPKLSPLQGRKNDQKIFPSH 580
QY 248 SSKMTTFPLKNEVSVKRVTPQKQMGKPLHLTLPLALQVYAGSGTLTALAKTKG 307
DB 561 SDKKYTFDQGLIIS-----NVEPGDEGYTCQITTKLMDVASSLTLLTC-DRPPP 631
QY 308 LHQENVLVWRATQLOKNTLCEWGP---TSPKMLSLKLENKAKVSKREK----- 356
DB 632 VHLQVTNAKHRYVTLN-----WTPGDNNSPILLEYVVEFDQMKENGHELRVAD 684
QY 357 -----PVW 359
DB 685 KGHVNLPLW 693
RESULT 61
ID IGA4_HAEIN STANDARD; PRT; 1849 AA.
AC P45386;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Immunoglobulin A1 protease precursor (EC 3.4.21.72) (IGA1 protease).
GN IGA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxId=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHIT HK61;
RX MEDLINE=92234949; PubMed=1373717;
RA Poulsen K., Reinhold J., Kilian M.;
RT "A comparative genetic study of serologically distinct Haemophilus
RT influenzae type 1 immunoglobulin A1 proteases."
RL J. Bacteriol. 174:2913-2921(1992).
CC -!- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A

```

```

CC PRODUCING INTACT FC AND FAB FRAGMENTS.
CC -!- CATALYTIC ACTIVITY: Cleavage of immunoglobulin A molecules at
CC certain Pro-Xaa bonds in the hinge region. No small molecule
CC substrates are known.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -!- SIMILARITY: Belongs to peptidase family S6.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M87491; AAA24968.1; -.
CC MEROPS: S06.001; -.
CC InterPro: IPR006315; Autotransport.
CC InterPro: IPR005546; Autotransporter.
CC InterPro: IPR000710; Peptidase S6.
CC InterPro: IPR004899; Peptactin.
CC Pfam: PF03797; Autotransporter; 1.
CC Pfam: PF0395; IGA1; 1.
CC PRINTS: PR00921; Peptactin; 1.
CC TIGRFAMs: TIGR01414; autotrans bar1; 1.
CC HydroLase; Serine protease; Transmembrane; Zymogen; Signal.
CC SIGNAL 1 25
CC CHAIN 26 1021 IMMUNOGLOBULIN A1 PROTEASE.
CC PROPEP 1022 1849 HELPER PEPTIDE (POTENTIAL).
CC ACT SITE 299 299 PROBABLE.
CC SEQUENCE 1849 AA; 202957 MW; 79A7D018C7150AEA CRC64;
Query Match 4.9%; Score 113.5; DB 1; Length 1849;
Best Local Similarity 19.1%; Pred. No. 9.1;
Matches 88; Conservative 58; Mismatches 135; Indels 179; Gaps 19;
QY 20 PAATQGNKVL-----GKKGDTVELTCTASQKSIQFHKN 55
DB 735 PPHFENNEVVEDDMINRNFATTMNTGMSLSYGRVANITSNITSNNAQVHIGKT 794
QY 56 SNOIKILNGQSFLLTKGPSKLNDRADSRSLMDQGNFPLIKNLKIEDSTYICEVEDQK 115
DB 795 GDTVCVRSDYGYVTCHNSNSEKA----- 819
QY 116 EEVQLVPLGLT-ANSDTHLLOGGSLTLTLSPPGSSPSVQCRSPRK-NIQG-KITLSVS 172
DB 820 -----LNSFNPTNLKGNVLTENASFTL-----GKNLPGTIOISIGTS 857
QY 173 QLELDQSGTWCTYLQKQKVEFKIDIVLAFOKASSIVYKKEGVEFSFPPLAFVEXL 232
DB 858 QVNLKENSIMHLTGNSVNVQNLNGLHGLAQAQDANKVTTYN-----TLTVNSL 907
QY 233 TGSGLMWAQBRASSKSWTF-DLKNEVSVKRVTPQKQMGKPLHLTLPLALQVY 291
DB 908 SONGSFFY-----WVDFTNKSKVYVVK----- 931
QY 232 AGSGNTLALAKTKG-LHQENVLV-WRATQLOKNTLCEWGPSPKMLSLKENKEA 349
DB 932 SATGNFTLQVADKGTGEPHNETLFDASNAT--RNMLEVTLANGSVRGAWKRYKLRVNG 989
QY 350 K-----VSRKREKRVWTLN-----PEAGMGCILSDGQVLLLESNIKVLPTWS 391
DB 990 RYDLNPEVEKRNQVDTTNTITTPNDIOADPSA-----QSNNEIARVE 1034
QY 392 TPVHPRASALPAPPTGSLPDPQTASALPDP-PASALP 430
DB 1035 TPVPP-----PAPATESAI-----ASEQPETRPAPETAQPA 1064

```

```

RESULT 62
VGRL_MOUSE
ID _VGRL_MOUSE STANDARD; PRT; 1333 AA.
AC P3569; OS05094; Q61517;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Vascular endothelial growth factor receptor 1 precursor (EC 2.7.1.112)
DE (VEGFR-1) (Tyrosine-protein kinase receptor FLT) (FLT-1) (Embryonic
DE receptor kinase 2).
DE FLT1 OR FLT OR EPRK2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CN NCBI_TaxID=10090;
RX [1]
RX SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Neonatal brain, and Placenta;
RC MEDLINE=93330572; PubMed=8393164;
RA Finerman Y., Kelleher K., Morris G.E., Bean K., Merberg D.M.,
RA Kriz R., Morris J.C., Sookdeo H., Turner K.J., Wood C.R.;
RT "Molecular cloning of murine FLT and FLT-1."
RL Oncogene 8:2293-2298(1993).
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=94181281; PubMed=8134130;
RA Choi K., Wall C., Hanratty R., Keller G.;
RT "Isolation of a gene encoding a novel receptor tyrosine kinase from
RT differentiated embryonic stem cells."
RL Oncogene 9:1261-1266(1994).
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6; TISSUE=Lung;
RC MEDLINE=98201626; PubMed=9524283;
RA Kondo K., Hirasaka S., Subbalekshmi E., Matsushima H., Shibuya M.;
RT "Genomic organization of the flt-1 gene encoding for vascular
RT endothelial growth factor (VEGF) receptor-1 suggests an intimate
RT evolutionary relationship between the 7-Ig and the 5-Ig tyrosine
RT kinase receptors."
RL Gene 208:297-305(1998).
CC -!- FUNCTION: RECEPTOR FOR VEGF AND PLGF. HAS A TYROSINE-PROTEIN
CC KINASE ACTIVITY. THE VEGF-KINASE LIGAND/RECEPTOR SIGNALING SYSTEM
CC PLAYS A KEY ROLE IN VASCULAR DEVELOPMENT AND REGULATION OF
CC VASCULAR PERMEABILITY.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases.
CC -!- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration-
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC
CC -----
CC EMBL; L07297; AAA40078.1; -
CC EMBL; X78568; CA5531.1; -
CC EMBL; D88689; BAA24498.1; -
CC PIR; I78875; I78875.
CC PIR; S49010; S49010.
CC HSSP; P17948; IGSV.
CC MGD; MGI:95558; Flt1.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_C2.
CC InterPro; IPR000719; Prot_Kinase.
CC InterPro; IPR001824; ReceptTykImIII.

```

| Qy | 33 | KGDTVELCTASOKKSIQIFHKWNSNQIKLGNQGSFLTTKGPSKLNDRAHSRSIMDQNF | 92 |
|-----------------------|-------|--|----|
| Query Match | 4.9% | Score 113; DB 1; Length 1333; | |
| Best Local Similarity | 19.4% | Pred. No. 6.4; | |
| Matches | 90; | Conservative 73; Mismatches 168; Indels 132; Gaps 21; | |

```

Db      570 EGEDKLKSCVANKFLYRDTM---ILLRTVNRNTHMHSISKQMATQDYSITL----- 620
QY      93 LIINKLKIEDSDTYICEV-----EDQKEEVOQLVFGLTANSDFHLQOGSLTTLTLESP 146
Db      621 NLIVINKVSLSDSGTACRARNIYTGEDILRTKEVLVRSEA---PHLLQNS---DYEVS 674
QY      147 PSSSPVQCRSPRG-----KN---IQ-----GGKTLVSQLELDQSGTWCT 185
Db      675 ISGSTTLDOCA-RGVPAPQITWFKNNHKIQDEPGIILGSGNSTLIERTEDEGVYCR 733
QY      186 VLQNGKVE-----FKIDIVLAFQKASSIY-----KKEGEQVEF 221
Db      734 ATNQGAVERSAAYLTIVQGTSDKSNLELTLTCTCVAATLFMLLTFLTRKLRSSSEVKT 793
QY      222 SF-----PLAFVYKLTGSGELM-WQAEPRASSSSSM-----ITFDLKN- 258
Db      794 DLSIIMDPDEVLEQCELRPYDASKWEFARERKLGSLRGAFGVQVQASAFGIKKS 853
QY      259 ---KEVSVRVTPDPLQWKKLPLHLTPQALPOVAGSGNTLTLALAKTGKLGHEVNLV 315
Db      854 PTCRTVAAYKMLGEGATASEYKALMTBLKTLTHGHILNVNLLGACTKQGGPL-----MV 908
QY      316 VWRATQLOKNLTCEVWGPTSPKMLSLKLENKAVSKREKPVWVLPNPAQMOCILSDS 375
Db      909 IVE-----YCK-YGNLSNYL-----KSKRDLFCINKDALHMLKKES 945
QY      376 GGVLESNTKVLPTWSTPHRPASALPAP-TGSLALPDPQTS 417
Db      946 LEFGLEQCGK-----PRLDSVSSSVTSSEFPEDRSVS 978

```

RESULT 63

LACH DROME STANDARD; PRT; 359 AA.

```

ID      LACH DROME STANDARD; PRT; 359 AA.
AC      Q24372;
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Lachesin precursor.
GN      LAC.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydrioidae; Drosophilidae; Drosophila.
OX      NCBI_Taxid=7227;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94038693; PubMed=8223276;
RA      Karlstrom R.O., Wilder L.P., Baetiani M.J.;
RT      "Lachesin: an immunoglobulin superfamily protein whose expression
RT      correlates with neurogenesis in grasshopper embryos.";
RL      Development 118:509-522(1993).
CC      -1- FUNCTION: MAY PLAY A ROLE IN EARLY NEURONAL DIFFERENTIATION AND
CC      AXON OUTGROWTH.
CC      -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
CC      similarity).
CC      -1- DEVELOPMENTAL STAGE: EXPRESSED ON DIFFERENTIATING NEURONAL CELLS
CC      FROM THE ONSET OF NEUROGENESIS IN BOTH THE CENTRAL AND PERIPHERAL
CC      NERVOUS SYSTEMS.
CC      -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC      -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; L13255; AAC37184.1; -.
CC      HSSP; P56276; TTKL.
CC      PDBase; Fbgno010238; Lac.

```

```

DR      InterPro; IPR007110; IG-like.
DR      InterPro; IPR003598; IG_c2.
DR      Pfam; PF00047; Ig_3.
DR      SMART; SM00408; IgC2; 2.
DR      PROSITE; PS50835; IG_LIKE; 3.
KM      Immunoglobulin domain; Cell adhesion; glycoprotein; GPI-anchor;
KM      Repeat; Signal; Lipoprotein.
FT      SIGNAL 1 25
FT      CHAIN 26 336
FT      PROPEP 337 359
FT      DOMAIN 29 130
FT      DOMAIN 135 221
FT      DOMAIN 226 317
FT      DISULFID 50 113
FT      DISULFID 157 204
FT      DISULFID 247 303
FT      CARBOHYD 92 92
FT      CARBOHYD 140 140
FT      LIPID 336 336
SQ      SEQUENCE 359 AA; 39998 MW; 513E2B9D80686 CRC64;

```

Query Match 4.8%; Score 112.5; DB 1; Length 359;
 Best local similarity 20.1%; Pred. No. 1.3;
 Matches 46; Conservative 39; Mismatches 109; Indels 35; Gaps 4;

```

QY      34 GDTVELCTASQKSIQPHMKNQIKILNGQSGFLTGPKNDRADRSRLMDQGNFP 93
Db      43 GGVTFPDCSVQYAEVYNVLFKTDSDPVFLTGSTLVTKDSRFSRLRDPNS---TYK 97
QY      94 LIINKLKIEDSDTYICEV-----VEDQKEEVOQLVFGLTANSDFHLQOGSLTTLTLESPG 148
Db      98 LQIKDIOETDAGTTCQVIVSTVHKVSEVLTSVRPVRVID-----NTQSVVASEG 150
QY      149 SSPSVQCRS-----PRGNIOGKTLVSQLELDQSGTWCTVLQNO 190
Db      151 SEVQMECVASGYPTTITWRRENNAILPTDSATVYGNLTIRIKSVKEDRGTYCVADNGV 210
QY      191 KVKERKIDIVLAFQKASSIYKKKEGEVESFPPLAFVYKLTGSGELM 239
Db      211 SKGDRNINVEVEFAPVITVPRPRIGQALQYDMDECHIEAYPPPAIVW 259

```

RESULT 64

C166 CHICK STANDARD; PRT; 588 AA.

```

ID      C166 CHICK STANDARD; PRT; 588 AA.
AC      P42292;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      CD166 antigen precursor (SCL glycoprotein) (BEN glycoprotein) (DM-
DE      GRASP protein) (UC7 protein).
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosaustra; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Gallus.
OX      NCBI_Taxid=9031;
RN      [1]
RP      SEQUENCE FROM N.A. AND SEQUENCE OF 34-53.
RP      TISSUE=Embryo;
RX      MEDLINE=92030150; PubMed=1931049;
RA      Tanaka H., Matsui T., Agata A., Tomura M., Kubota I.,
RA      McFarland K.C., Kohr B., Lee A., Phillips H.S., Shelton D.L.;
RT      "Molecular cloning and expression of a novel adhesion molecule, SCL";
RL      Neuron 7:535-545(1991).
CC      [2]
CC      SEQUENCE FROM N.A.
CC      MEDLINE=91337449; PubMed=1873027;
CC      Burns F.R., von Kamen S., Guy L., Raper J.A., Kamholz J.,
CC      Chang S.;
CC      "DM-GRASP, a novel immunoglobulin superfamily axonal surface protein
CC      that supports neurite extension.";
CC      Neuron 7:209-220(1991).
CC      [3]

```

| | |
|---|---|
| R | SEQUENCE FROM N.A., AND SEQUENCE OF 34-52; 87-93; 504-512 AND 569-582. |
| R | TISSUE-Bursa of fabricius; |
| R | MEDLINE=92302224; PubMed=1608932; |
| R | Pourquie O., Cortel C., le Caet J.-P., Rossier J., le Douarin N.M.; |
| R | "BNR, a surface glycoprotein of the immunoglobulin superfamily, is expressed in a variety of developing systems."; |
| R | Proc. Natl. Acad. Sci. U.S.A. 89:5261-5265(1992). |
| R | [4] |
| R | POSSIBLE FUNCTION. |
| R | MEDLINE=92211411; PubMed=113497; |
| R | Pourquie O., Hallonet M.E.R., le Douarin N.M.; |
| R | "Association of BEN glycoprotein expression with climbing fiber axonogenesis in the avian cerebellum."; |
| R | J. Neurosci. 12:1548-1557(1992). |
| R | - FUNCTION: HOMOPHILIC ADHESION MOLECULE ASSOCIATED WITH CLIMBING FIBER AXONGENESIS. SUPPORTS NEURITE EXTENSION. |
| R | - SUBCELLULAR LOCATION: Type I membrane protein. |
| R | - TISSUE SPECIFICITY: EXPRESSED IN A RESTRICTED POPULATION OF AXONS. WITHIN THE SPINAL CORD IT IS LOCALIZED TO AXONS IN THE DORSAL PUNICULUS, MIDLINE FLOOR PLATE CELLS, AND MOTONEURONS. FOUND IN EPITHELIA, CENTRAL NERVOUS SYSTEM, AND HEMOPOIETIC CELLS. EACH DEVELOPING SYSTEM CARRIES DIFFERENTLY GLYCOSYLATED MOLECULAR FORMS OF BEN. |
| R | - DEVELOPMENTAL STAGE: WIDELY EXPRESSED DURING EMBRYONIC DEVELOPMENT. |
| R | - SIMILARITY: Contains 3 immunoglobulin-like C2-type domains. |
| R | - SIMILARITY: Contains 2 immunoglobulin-like V-type domains. |
| R | ----- |
| R | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). |
| R | ----- |
| R | EMBL; S63276; AAB20170.1; - |
| R | EMBL; M76678; AAA48602.1; - |
| R | EMBL; X64301; CAA45579.1; - |
| R | PIR; A45254; A45254. |
| R | PIR; JH0464; JH0464. |
| R | PIR; JH0506; JH0506. |
| R | HSSP; Q13740; JKJC. |
| R | InterPro; IPRO07110; Ig-I-like. |
| R | InterPro; IPRO03599; Ig. |
| R | InterPro; IPRO03006; Ig_MHC. |
| R | Pfam; PF00047; Ig_5. |
| R | SMART; SM00409; IG_3. |
| R | PROSITE; PS50835; IG_LIKE; 4. |
| R | PROSITE; PS00290; Ig_MHC; FALSE NEG. |
| R | KW Cell adhesion; Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal. |
| R | FT SIGNAL 1 33 |
| R | FT CHAIN 34 588 CD16 ANTIGEN. |
| R | FT DOMAIN 34 532 EXTRACELLULAR (POTENTIAL). |
| R | FT TRANSMEM 533 553 POTENTIAL. |
| R | FT CYTOPLASMIC (POTENTIAL). |
| R | FT DOMAIN 554 588 IG-LIKE V-TYPE 1. |
| R | FT DOMAIN 42 126 IG-LIKE V-TYPE 2. |
| R | FT DOMAIN 131 240 IG-LIKE C2-TYPE 1. |
| R | FT DOMAIN 251 333 IG-LIKE C2-TYPE 2. |
| R | FT DOMAIN 338 414 IG-LIKE C2-TYPE 3. |
| R | FT DOMAIN 421 501 IG-LIKE C2-TYPE 3. |
| R | FT DISULFID 49 119 POTENTIAL. |
| R | FT DISULFID 163 226 POTENTIAL. |
| R | FT DISULFID 276 319 POTENTIAL. |
| R | FT DISULFID 359 397 POTENTIAL. |
| R | FT DISULFID 440 490 POTENTIAL. |
| R | FT CARBOHYD 101 101 N-LINKED (GLCNAC . .) (POTENTIAL). |
| R | FT CARBOHYD 173 173 N-LINKED (GLCNAC . .) (POTENTIAL). |
| R | FT CARBOHYD 199 199 N-LINKED (GLCNAC . .) (POTENTIAL). |
| R | FT CARBOHYD 271 271 N-LINKED (GLCNAC . .) (POTENTIAL). |
| R | FT CARBOHYD 312 312 N-LINKED (GLCNAC . .) (POTENTIAL). |
| R | FT CARBOHYD 366 366 N-LINKED (GLCNAC . .) (POTENTIAL). |

Query Match 8 4.8%; Score 112; DB 1; Length 588;
Best Local Similarity 19.2%; Pred. No. 2.6;
Matches 91; Conservative 66; Mismatches 175; Indels 140; Gaps 19;

| | | | | |
|----|----------|---------|-----------|-------------------------------------|
| FT | CARBOHYD | 462 | 462 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 485 | 485 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 504 | 504 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CONFLICT | 1 | 10 | MMEPPAAR -> MEPPSRRP (IN REF. 3). |
| FT | CONFLICT | 25 | 25 | A -> S (IN REF. 3). |
| FT | CONFLICT | 112 | 113 | SD -> RH (IN REF. 3). |
| FT | CONFLICT | 329 | 329 | A -> T (IN REF. 2). |
| FT | CONFLICT | 401 | 402 | LQ -> HK (IN REF. 2). |
| SO | SEQUENCE | 588 AA; | 65726 MW; | 2A28612D0164531E CRC64; |

Query Match 8 4.8%; Score 112; DB 1; Length 588;
Best Local Similarity 19.2%; Pred. No. 2.6;
Matches 91; Conservative 66; Mismatches 175; Indels 140; Gaps 19;

| | |
|----|---|
| QY | RHLHLVQLALLPATGKRVGLGKKGDPVELETCTASQKSF-HWK---NSNQIKL 62 |
| DB | 16 RPLICLLALALCPMPALGLYTVANVAVYDITIMPCRLBEPDGLMGKKYKEMPNSPFFIA 75 |
| QY | 63 GNQGSFLTKGPSKLANDADRRSLMDQGNPLIIKLIKIDSDPTIYCEV--EDQKEVOL 120 |
| DB | 76 PRSS--TKKNVQYDVPDYDKDRSLSENENTLISKARISDEKRFVCMLVTEDEVSEPT 132 |
| QY | 121 LVFGLTANSDPHLLQGSLLFTLTSPGSSPSVQCSRPG-----KN-----IQGSKTL 169 |
| DB | 133 VVKFKPQSPQPEILHQDPLETEKRLKGLGECVYARDSPEGNVTMYKNGRVLQPVAEVVI 192 |
| QY | 170 SVSQLELDQSG-----TWCTTVL-----QNKQVEF----- 195 |
| DB | 193 NLRKVENRSGLFTWTSLSQYPTKEDANAKFTCIYVHYGSPGQKTIQSPFVVDVHYPT 252 |
| QY | 196 -KIDIVVLAQKASIIYKKEGGEVSEFPLAFVEKLTSGE-----LW---QAKRA 245 |
| DB | 253 EKVIRPLLS--QSSTI---KEGNVTL-----KSGNGNPPQGFLEFIYIGEFGI 298 |
| QY | 246 SSSSWITPDLKKNKEVSKRVTDPK-----LQNGKPLH 281 |
| DB | 299 RSSSTYWTDRNRATGEYKCSLIDSMMDATITVHYLDLQLPSPSEVTKQIGEALPVS 358 |
| QY | 282 LTLPEALP-----QYAGSGNL---TLALEAKTGLHDEVNLV- 316 |
| DB | 359 CTISSSNATVFWIKDTRKKTSPFSLYQDAGNVICGTTIQEYVGLKKRKLKIVE 418 |
| QY | 317 -----MRATQLOKMLTCEYWGTSFCLMLSLKLENKEAKVSKREKPVVY 360 |
| DB | 419 GKPIKMTKTKTNTNKKSKTIYCHVEG--FFKPAVQWTVTSGSLINKTEETKYV 470 |

RESULT 65
IL1S_CERAE STANDARD; PRT; 393 AA.

IL1S_CERAE AC Q29612;
15-JUL-1999 (Ref. 38, Created)
15-JUL-1999 (Ref. 38, Last sequence update)
10-OCT-2003 (Ref. 42, Last annotation update)
DE Interleukin-1 receptor, type II precursor (IL-1R-2) (IL-1R-beta).
GN IL1R2 OR IL1RB.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Cercopithecus.
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
[1]
SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT), AND SEQUENCE OF 125-145.
MEDLINE=96355446; PubMed=8702856.
RX Liu C., Hart R.P., Liu X.J., Clevenger W., Maki R.A., Souza E.B.;
RA "Cloning and characterization of an alternatively processed human
RT type II interleukin-1 receptor mRNA."
RL J. Biol. Chem. 271:20965-20972(1996).
CC -1- FUNCTION: RECEPTOR FOR INTERLEUKIN-1 ALPHA (IL-1A), BETA (IL-1B),
CC AND INTERLEUKIN-1 RECEPTOR ANTAGONIST PROTEIN (IL-1RA).
CC -1- SUBCELLULAR LOCATION: THE LONG ISOFORM IS A TYPE I MEMBRANE
CC PROTEIN WHILE THE SHORT ISOFORM IS SOLUBLE.
CC -1- ALTERNATIVE PRODUCTS:

```
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q29612-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q29612-2; Sequence=VSP_002663;
CC -1 SIMILARITY: Belongs to the interleukin-1 receptor family.
CC -1 SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U64092; AAB05876.1; -.
CC EMBL; U64093; AAB05877.1; -.
CC HSRP; P14778; 11TB.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003599; IG.
CC InterPro; IPR004074; IL1_receptor/IL1.
CC InterPro; IPR004077; IL1_receptor/IL1.
CC Pfam; PF00047; Ig; 3.
CC PRINTS; PRO1539; INTERLEUKINR2.
CC PRINTS; PRO1536; INTERLEUKINR12F.
CC SMART; SM00409; IG; 3.
CC PROSITE; PS50835; IG_LIKE; 3.
CC Immunoglobulin domain; Receptor; Glycoprotein; Transmembrane; Signal;
CC Repeat; Alternative splicing.
CC FT CHAIN 1 13 POTENTIAL.
CC FT SIGNAL 1 13 INTERLEUKIN-1 RECEPTOR, TYPE II.
CC FT DOMAIN 14 347 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 348 368 POTENTIAL.
CC FT DOMAIN 29 120 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 134 221 IG-LIKE C2-TYPE 1.
CC FT DOMAIN 237 342 IG-LIKE C2-TYPE 2.
CC FT DISULFID 50 108 IG-LIKE C2-TYPE 3.
CC FT DISULFID 152 207 BY SIMILARITY.
CC FT DISULFID 258 326 BY SIMILARITY.
CC FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT VARSPLIC 297 393 Missing (in isoform Short).
CC FT FTId=VSP_002663.
CC SQ SEQUENCE 393 AA; 44778 MW; DAD746C5DE59A8 CRC64;

Query Match 4.8%; Score 111.5; DB 1; Length 393;
Best Local Similarity 20.3%; Pred. No. 1.7;
Matches 70; Conservative 52; Mismatches 120; Indels 103; Gaps 14;

Oy 7 FRHLVLVLA---LLPATOG---NKVVLGK-----KGDYVELACT-----ASQ 45
Db 2 FRLLVLVGVSAFTIQPAHATGAASCPYRHHYKREPRLEBEPALACPOVPYVLAHSV 61
Oy 46 KKSIOFHNKSNQIKILGNOSFLTKGPSKLNDRADSRSLWDQGNFPLIKNLKIEDSD 105
Db 62 SPHNLVTHKNDASAMV-----PGEETR-----WMAQGALMLPALQ-EDSG 104
Oy 106 TYICGVED---QKEVQLVFLGLTANSDTHLLOQSITL-----TLSEPP 147
Db 105 TYICTTRNAYSCKKSIELRPFENDASLPISYDQIITLSTFGVLVCPDLREFTRDKTD 164
Oy 148 GSPSPVOCSPRGK-----NIQGGKTLVSQLEPDQSGTWTCTV-----186
Db 165 GKIQWYKQFLPDKNKEKELSVRGTHLLVHVALIEDAGYRVCVLTFAHGGQYITRNI 224
Oy 187 -LQNGKYVEFKIDIVLAFQKASSIVYKKGEGVQFSPPLAFTVEKLTGSGE----LWV 240
Db 225 ELRIKKKEETIPVILSPKLTISASLSGR-----LTIPCKVFPLOTGTPPLTTLMMW 274
```

```
Oy 241 QAERASSKSWITFDLKNKEVSVKRYQDPKLQWKKLPLHLTP 285
Db 275 TANDTHESAY-----PGGRVTEGPRQETSENNENYIEVP 309

RESULT 66
SCA4_RICFE
ID_SCA4_RICFE STANDARD; PRT; 981 AA.
AC Q9AJ37;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Antigenic heat-stable 120 kDa protein (PS120) (120 kDa antigen)
DE (Protein PS 120) (Fragment).
GN SCA4 OR D.
OS Rickettsia felis (Rickettsia azadi).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=42862;
RN [1]
RA SEQUENCE FROM N.A.
RA Sekeyova Z., Roux V., Raoult D.;
RT "Phylogenetic analysis of Rickettsia spp. by comparing the sequence of
RT gene D coding for an intracytoplasmic protein.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1 SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF196973; AAK1304.1; -.
CC Antigen.
CC FT NON_TER 1 1
CC FT NON_TER 981 981
CC SQ SEQUENCE 981 AA; 107514 MW; 7F18F421E2C262E1 CRC64;

Query Match 4.8%; Score 111.5; DB 1; Length 981;
Best Local Similarity 24.4%; Pred. No. 5.5;
Matches 88; Conservative 48; Mismatches 150; Indels 75; Gaps 19;

Oy 110 EVEDQKEVQLVFLGLTANSD--THLQGGSLTLLTS-PRGSSPSVQCSPRGKNTQGG 166
Db 103 EIEBKDKTLISAFPGNPANREFIDKALENPELKKLSIEIAGYKNVLSTYSANGYGGG 162
Oy 167 -----KTLVSQLEL---QDSGTWTCTYVLOQKYEFTIDIVLAFQKASSIVYKKE 215
Db 163 FKPYQWENQISASPLRATVVVNDAGDELCTV--NETTYKTK-----PFTVAKOD 209
Oy 216 GEQV-----EFSPPLAFTVEKLTGSGELMWQARASSKSWITFDLKNKEVSVKRYQD 269
Db 210 GTQVQINSYREIDPPI--KDKKADSGMILSNVALKAQSTK-----PSKQKAYFTTAY 260
Oy 270 PKLQWKKLPLHLTPALPOYAGSGNLTALAEKTKLHDEQVNLVVMARATLOKNTLCE 329
Db 261 EESGNGKQKQLEISPKRL-KFADGDPAAVYIEHGGIY---TLAVTRGK--YKEMRE 314
Oy 330 VWGFTSPKMLSLKLENKEAVYSR--EKPVWVNLNPEAGMWOCLISDSGOVLLESNIKVL 387
Db 315 VELNQGGSVDSQRTAEDLTQVQGSQETPQPIITPQOE-----LKSISIT- 360
Oy 388 PTWSPVYPRASA-LPAAPPTGSAALPDQGTASALDPP-----AASALPAALAVISFLGIG 442
Db 361 PT-TTQVPPITPAQNPLOPETSQMPQOVN-----PMLVATYATLSTSMODLNTYVAG 414
Oy 443 L 443
Db 415 L 415
```

RESULT 67
 FPRP_MOUSE STANDARD; PRT; 879 AA.
 AC Q9WV91;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Prostaglandin F2 receptor negative regulator (Prostaglandin
 F2-alpha receptor regulatory protein) (Prostaglandin F2-alpha receptor
 associated protein).
 DE associated protein).
 GN PIGRN OR FPRP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Meng L., Falla N., Van Den Heuvel R., Van Vlaeselaer P.,
 RA Raymackers J., Lowik C., Merregaert J.;
 RT "The monoclonal antibodies 18D7/91F2 recognize a receptor regulatory
 protein on mouse bone marrow stromal cells";
 RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Inhibits the binding of prostaglandin F2-alpha (PGF2-
 alpha) to its specific FP receptor, by decreasing the receptor
 number rather than the affinity constant. Functional coupling with
 the prostaglandin F2-alpha receptor seems to occur (By
 similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
 reticulum and trans-Golgi network (By similarity).
 CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL, AF152344; AAD38383.1; -
 DR MGD; MGI:1277114; Pctfzn.
 DR GO; GO:0005515; F:protein binding; ISS.
 DR GO; GO:0017148; P:negative regulation of protein biosynthesis; ISS.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003596; IG_v.
 DR Pfam; PF00047; Ig; 6.
 DR SMART; SM00406; IGv; 1.
 DR PROSITE; PS50835; IG LIKE; 5.
 DR Signal; Immunoglobulin domain; Transmembrane; Glycoprotein; Repeat.
 KW SIGNAL
 FT SIGNAL 1 21
 FT CHAIN 22 879
 FT
 FT DOMAIN 22 832
 FT TRANSSEM 833 853
 FT DOMAIN 854 879
 FT DOMAIN 22 137
 FT DOMAIN 149 263
 FT DOMAIN 276 389
 FT DOMAIN 406 536
 FT DOMAIN 544 662
 FT DOMAIN 688 813
 FT DOMAIN 424 427
 FT SITE 89 91
 FT SITE 703 705
 FT DISULFID 43 119
 FT DISULFID 169 247
 FT DISULFID 299 373
 FT DISULFID 429 515
 FT DISULFID 571 655
 FT DISULFID 711 793
 FT CARBOHYD 44 44
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 POTENTIAL.
 REGULATOR.
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 IG-LIKE C2-TYPE 1.
 IG-LIKE C2-TYPE 2.
 IG-LIKE C2-TYPE 3.
 IG-LIKE C2-TYPE 4.
 IG-LIKE C2-TYPE 5.
 IG-LIKE C2-TYPE 6.
 ENDOLASMIC RETICULAR RETENTION SIGNAL.
 CELL ATTACHMENT SITE (POTENTIAL).
 CELL ATTACHMENT SITE (POTENTIAL).
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 POTENTIAL).

FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 383 383 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 525 525 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 600 600 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 691 691 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 879 AA; 98707 MW; 0D3D3EB271478D92 CRC64;
 Query Match 4.8%; Score 110.5; DB 1; Length 879;
 Best Local Similarity 19.5%; Pred. No. 5.6; 203; Indels 125; Gaps 21;
 Matches 100; Conservative 86; Mismatches
 QY 14 LQALIPAAATGKGNKLVGKGGDTVELTCTASQ-----KKSIQFMKNSNQIKLG---- 63
 Db 279 LQLAV-----PRTSVTEGKDLDSCTTTTRVDVDRVRYVMYFKKTDSTLASHML 331
 QY 64 ---NQGSFLTKGPKSLNDRADRSRLMDQGNFPLIIKNIKIEDSDYICV-----ED 113
 Db 332 ARLDRLSVHSPHVALSHVDR-----SYHLVADVSKENSGYLCVALMAPGHR 384
 QY 114 QKEEVLAVGLTANSDTHLQGSULTLESPPGS---PSVQCRSPRGKNIQCGKTL 170
 Db 385 SMKIVAMSPGVSQVTLPEPYQVYLNASKVPGFSDDPTELOCRVITDKLEAGVRLT 444
 QY 171 V-----SOLEQDSGTWTCTVLQOKKVEFKIDIVLAFQKA--SSIVYK 213
 Db 445 VSMYRTRNDVDVASELLAVMDGDTLRYGSRK-----QRAQDGEIF 491
 QY 214 KEGBOVEFSPPLAFYVEKLTGSGELMWMQERASSKSMI--TFPLKKNKSVKRYTDPKL 272
 Db 492 KEHTDTR--FNRIORTTEEDBNYCVVASAWARQNNMSVSKDFSPVNI FNASEBSVL 550
 QY 273 QMGKKPLHLTLQALPOVAGSNLTLALFAKTKLHQEVNLVVMRAVTOLOKULCEVWG 332
 Db 551 VVAKRP-----KPFPAAGTFEWTCKVSSKNIKSPYSLITKEKVGDIS---S 598
 QY 333 PTPSKMLTSL-----KLEN-----KEAKVSREKRPVWLNPEAGMOCCL- 372
 Db 599 PNEKTYIISLDQSVKLEWMTDASRDVGLKEVQDEDERMYMYQODSDAGLYRCMVT 658
 QY 373 --SDSGCVL-----LESNIKV---LPTWSTPVHPRASALPAPGTGALP--DPQT 415
 Db 659 AWPPIGGSLRREKATSIENPIEIDPQTSGPTFNASVH---SDTPSVYRGDLKLCFIVTV 715
 QY 416 ASALPDPPEASALPALAVISFLGIGLVACTL 449
 Db 716 EGAVLDPDDMAFDVSMFAVHSF---GLDKAPVL 745
 RESULT 68
 RTN1_RAT STANDARD; PRT; 777 AA.
 ID RTN1_RAT
 AC Q64546; Q64547;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Reticulon 1 (Neuroendocrine-specific protein) (S-rex).
 GN RTN1 OR NSP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS RTN1-B AND RTN1-S).
 RC STRAIN=Wistar; TISSUE=Brain cortex;
 RX MEDLINE=96386034; PubMed=8793864;
 RA Baka I.D., Ninkina N.N., Pinon L.G.P., Adu J., Davies A.M.,
 RA Georgiev G.P., Buchman V.L.;
 RT "intracellular compartmentalization of two differentially spliced s-
 rex/NSP mRNAs in neurons";
 RL Mol. Cell. Neurosci. 7:289-303 (1996).
 CC -1- FUNCTION: May be involved in neuroendocrine secretion or in

```

CC membrane trafficking in neuroendocrine cells.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=RTN1-B; Synonyms=S-RexB;
CC IsoId=O64548-1; Sequence=Displayed;
CC Name=RTN1-S; Synonyms=S-RexS;
CC IsoId=O64548-2; Sequence=VSP_005647, VSP_005648.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN CENTRAL AND
CC PERIPHERAL NERVOUS SYSTEM OF NEWBORN AND ADULT RATS. LOW LEVELS
CC HAVE BEEN ALSO DETECTED IN HEART, ADRENAL GLAND AND SPLEEN.
CC EXPRESSION OF ISOFORM RTN1-B IS RESTRICTED TO PARTICULAR NEURONAL
CC TYPES.
CC -1- DEVELOPMENTAL STAGE: DETECTED ON EMBRYONIC DAY E10 IN THE
CC HINDBRRAIN AND IN E11 IN THE FOREBRAIN. DURING THE NEXT 3 EMBRYONIC
CC DAYS THE LEVELS OF S-REXS INCREASES AND REMAINS STABLE AT E13 IN
CC THE HINDBRRAIN AND AT E14 IN THE FOREBRAIN. THE LEVELS OF S-REXB
CC DOES NOT CHANGE AS SIGNIFICANTLY DURING DEVELOPMENT OF THE
CC HINDBRRAIN.
CC -1- SIMILARITY: Contains 1 reticulation domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL, U17604; AAC3046.1; -
CC DR InterPro, IPR003388; Reticulon.
CC DR Pfam, PF02453; Reticulon; 1.
CC DR PROSITE, PS50845; RETICULON; 1.
CC KM Endoplasmic reticulum; Alternative splicing; Transmembrane.
CC FT TRANSMEM 604 624 POTENTIAL.
CC FT DOMAIN 550 777 POTENTIAL.
CC FT POLY-LEU.
CC FT VARSPLIC 1 569 MISSING (in isoform RTN1-S).
CC FT VARSPLIC 570 589 /FTID=VSP_005647.
CC FT IGPISGDIPLPPLPFPNCKK -> MQATADSTKMDQWSNM
CC FT KSQ (in isoform RTN1-S).
CC FT /FTID=VSP_005648.
CC SQ SEQUENCE 777 AA; 83001 MW; AF7479C5DF2BD0AC CRC64;
Query Match 4.7%; Score 110; DB 1; Length 777;
Beat Local Similarity 19.4%; Pred. No. 5.1;
Matches 77; Conservative 50; Mismatches 142; Indels 128; Gaps 15;
QY 95 IIKNLIKIEDSD--TYICEVEDEKKEVOLLV-----FGLTANSDFHLLQGSLVTLT 143
DB 247 LIKDNLFESTAPYIDELSDQHRMSLVTAIVKTLTFEIGPVMATATHTETPEKODCL 306
QY 144 EEPSSSPSVQCRSPRGKNIQGGKILSVSQLELDOSGTCTVQLQNKKNVEKIIIVILA 203
DB 307 KSPSPVPTVTVSEPEDD-----SPGSVTPPSSGT-EPBAASQGGSVSEDELLAA 357
QY 204 FOKASSIIVKKEGEVFEFPLAFVTEKLTSGELMWMQBARSSSKSMVTLTKKKEYSV 263
DB 358 IYEAGKLSI---ETTESRPVGOADR----- 381
QY 264 KRVTDPKLQMGKKLPLHLTPQALPQVYAGSNLTLEAKTGKLEHENVLVMRATOLQ 323
DB 382 -----PKVKARSGRP---TTPSSLDQKAS-----AESG--DSEIELVSEDPMAE 422
QY 324 KNLTC-----EYWG-----PTSPKMLSLKENKEAKVS-----KRE 355
DB 423 DALPSGYVSFGHVSQPPSPASPSIQYSTILREERELELDSELIISCDASSASESPKRE 482
QY 356 KPVWVNLNPAGMOCCLSDSGVLLSNNIKVLPTWSTPHRPAALPAPPTGSALPDPTOT 415

```

```

DB 483 QDSPPMKP--GVLDALIREETSSRATEER---APSGHGVPEBDDPILSFTPTVLQGRPESS 537
QY 416 ASALPDP-----PAAALPAA 432
DB 538 GDGAPVPEPPKXQQQKPEEAHVSSSQSPAAETIGPL 574
RESULT 69
GCC2_MOUSE STANDARD, PRT, 1679 AA.
AC Q8CHG3; Q8BR44; Q8R205; Q9CT45;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE GRP and coiled-coil domain-containing protein 2 (Golgi coiled coil
DE protein GCC185).
GN GCC2 OR KIA0336.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OK NCBI_TaxId=10090;
RN [1]
RP SEQUENCE OF 1-769 FROM N.A.
RC TISSUE=Breast cancer;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Scheet C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schenker C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uediri T.B., Toshitsuki S., Carranti P., Prange C.,
RA Rana S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Boeck S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Hellon E., Kettleman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schin J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE OF 1-442 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain; and Embryo;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nishida I., Oshio N., Saito R., Suzuki H., Yamana K., Kiyosawa H.,
RA Yagi K., Tomaru Y., Haegawa Y., Nogami A., Schonbach C., Gotohori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schmitt L.M., Kanapin A., Matasov H., Batilov S., Beisel K.W.,
RA Blake J.A., Brad D., Brusic V., Chochia S., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.P., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustinch S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.A.,
RA Kongway A., Kurichin I.V., Lee Y., Lehman R.D., Lyons P.A.,
RA Maggiori D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nunata K., Okido T., Pavan M.J., Petrea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Warneke Y., Wells C.,
RA Wilting L.G., Wyszewski A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Zimin P., Hayatsu N.,
RA Hitozane-Tsukikawa T., Kono H., Nakamura M., Sakazune N., Sato K.,
RA Shitaka T., Waki K., Kawai J., Aizawa K., Arawaka T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,

```

RA Yaenishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RT Nature 420:563-573(2002).
 RN [3]
 RN SEQUENCE OF 49-1679 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22579291; PubMed=12693553;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
 RA Nakajima D., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries";
 RL DNA Res. 10:35-48(2003).
 CC -1- FUNCTION: Probably involved in maintaining Golgi structure (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; peripheral membrane protein
 CC associated with the trans-Golgi network (By similarity).
 CC -1- DOMAIN: Extended rod-like protein with coiled-coil domains.
 CC -1- SIMILARITY: Belongs to the golgin family.
 CC -1- SIMILARITY: Contains 1 GRP domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 DR EMBL; BC027339; AA027339.1; ALT_INIT.
 DR EMBL; AK011206; BA027466.2; -
 DR EMBL; AK045701; BAC32463.1; -
 DR EMBL; AB093232; BAC41416.1; -
 DR PIR; P70649; P70712.
 DR MGD; MGI:1917547; 2600014C01R1K.
 DR InterPro; IPR000237; GRIP domain.
 DR InterPro; IPR002017; Spectrin.
 DR Pfam; PF01465; GRIP; 1.
 DR PROSITE; PS50913; GRIP; 1.
 DR KMW Golgi stack; Coiled coil.
 FT DOMAIN 31 1613 COILED COIL (POTENTIAL).
 FT DOMAIN 83 501 GLU-RICH.
 FT DOMAIN 1604 1654 GRP.
 FT CONFLICT 440 440 I->M (IN REF. 2).
 SQ SEQUENCE 1679 AA; 194443 MW; 6E2F84263E5E0103 CRC64;
 Query Match 4.7%; Score 110; DB 1; Length 1679;
 Best Local Similarity 20.0%; Pred. No. 14;
 Matches 87; Conservative 67; Mismatches 133; Indels 148; Gaps 18;
 QY 15 QALPAPATQGNKVVLGKGDPTVELTCTASOKSIQ--FWKNSNQIKITGNGSFLTK 71
 DB ELRLTVLRVSEBEKILIRK---ELDAVTSAKELQDLLEMKVTNEKSLDNG--TLST 871
 QY 72 GPSKLNDRADSRSRSLMDQ-----GNPLTIKN-----LK-----IE 102
 DB 872 QVEBLSQTLHSRNEVHDEKVLVIHEHNRLLTKQRESELQVRAELITLKSLSEKSPSVK 931
 QY 103 DSDTYICEVEDQKE-----EVLVLFGLTANSDTHLQGOQSLTTLTLESPFG 148
 DB 932 DQLSLVLEKEIKELSEKSKDKDEKISKILVAVKAKKELDSNRKQGTRELELS--- 988
 QY 149 SPSVQCSPPGKXNIQGGKITSVSGLE-LQDSGWTCTVLOKVKVEKIDIVLAPKA 207
 DB 989 -----VASEKDR-----LSASKVEFLQCAESYKSLLE----- 1016
 QY 208 SSIYKKEGGEVSEFPLAFVTEKLTGSGELMWQAEKSSKSMI--TFPLKKNKEVSVK 264
 DB 1017 ----YDKOSEQD-----VEK-----ERAHNFERHIEDLTQKLRNSTCOYE 1053

QY 265 RVTDPRK-----LQMGKCLPLHLTPGALPOYAGSGNLTALAKTKLHQEVNLV 316
 DB 1054 RLTSNDNEDLARIETTLQANK-----LLEAQILLEVQKXGVVEKELDAEBLQKEKXKEHV 1109
 QY 317 MRATQLOKNTLCEVWGTSFKMLSLKLENKEAK-----VSKREKPVWVLPENKGM 368
 DB 1110 STVNELE-----ELQLQFQEKKKQDKTMQEBELVYKDAQQTTLNMMEIADY 1156
 QY 369 QCLSDSGVLELESN 383
 DB 1157 ERLMKELNQKLTNKN 1171
 RESULT 70
 UN89 CAEEL STANDARD; PRT; 6632 AA.
 ID UN89 CAEEL
 AC 001761; 017362;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Muscle M-line assembly protein unc-89 (Uncoordinated protein 89).
 GN UNC-89 OR C09D1.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RC STRAIN=Br1sc01 N2;
 RX MEDLINE=96180278; PubMed=8603916;
 RA Benian G.M., Tinley T.L., Tang X., Borodovsky M.;
 RT "The Caenorhabditis elegans gene unc-89, required for muscle M-line
 RT assembly, encodes a giant modular protein composed of Ig and signal
 RT transduction domains";
 RL J. Cell Biol. 132:835-848(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Br1sc01 N2;
 RA Du Z., Le T.T., Wilson R.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Structural component of the muscle M-line. Myofibrillar
 CC lattice assembly begins with positional cues laid down in the
 CC basement membrane and muscle cell membrane. UNC-89 responds to
 CC these signals, localizes, and then participates in assembling an
 CC M-line.
 CC -1- TISSUE SPECIFICITY: Localizes to the middle of A-bands.
 CC -1- SIMILARITY: Contains 1 DBL-homology (DH) domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type III domain.
 CC -1- SIMILARITY: Contains 49 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 PH domain.
 CC -1- SIMILARITY: Contains 5 RGD domain.
 CC -1- SIMILARITY: Contains 1 SH3 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 DR EMBL; U33058; AA00542.1; -
 DR EMBL; AF003131; AB554132.2; -
 DR PDB; 1FHO; 20-DEC-00.
 DR WormPep; C09D1.1; CE30426.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR007110; Ig_II.
 DR InterPro; IPR003598; Ig_c2.

DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001849; PH_
DR InterPro; IPR007850; RCSD.
DR InterPro; IPR00219; RhGEF.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; Ig; 47.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF05177; RCSD; 5.
DR Pfam; PF00621; RhGEF; 1.
DR Pfam; PF00018; SH3; 1.
DR SMART; SM00408; IGC2; 23.
DR SMART; SM00325; RhGEF; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50010; DH_2; 1.
DR PROSITE; PS50835; IG_LIKE; 49.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50002; SH3; 1.
KW Muscle protein; Immunoglobulin domain; Repeat; SH3 domain;
3D-structure.
FT DOMAIN 63 127 SH3.
FT DOMAIN 152 330 DH.
FT DOMAIN 342 498 PH.
FT DOMAIN 547 633 IG-LIKE C2-TYPE 1.
FT DOMAIN 648 736 IG-LIKE C2-TYPE 2.
FT DOMAIN 748 838 IG-LIKE C2-TYPE 3.
FT DOMAIN 946 1033 IG-LIKE C2-TYPE 4.
FT DOMAIN 1044 1132 IG-LIKE C2-TYPE 5.
FT DOMAIN 1140 1227 IG-LIKE C2-TYPE 6.
FT DOMAIN 1272 1315 THR-RICH.
FT DOMAIN 1375 1475 RCSD 1.
FT DOMAIN 1479 1585 RCSD 2.
FT DOMAIN 1597 1695 RCSD 3.
FT DOMAIN 1700 1799 RCSD 4.
FT DOMAIN 1800 1860 RCSD 5.
FT DOMAIN 1982 2067 IG-LIKE C2-TYPE 7.
FT DOMAIN 2071 2163 IG-LIKE C2-TYPE 8.
FT DOMAIN 2171 2261 IG-LIKE C2-TYPE 9.
FT DOMAIN 2269 2359 IG-LIKE C2-TYPE 10.
FT DOMAIN 2367 2455 IG-LIKE C2-TYPE 11.
FT DOMAIN 2463 2564 IG-LIKE C2-TYPE 12.
FT DOMAIN 2563 2651 IG-LIKE C2-TYPE 13.
FT DOMAIN 2657 2746 IG-LIKE C2-TYPE 14.
FT DOMAIN 2754 2858 IG-LIKE C2-TYPE 15.
FT DOMAIN 2887 2980 IG-LIKE C2-TYPE 16.
FT DOMAIN 2994 3081 IG-LIKE C2-TYPE 17.
FT DOMAIN 3087 3183 IG-LIKE C2-TYPE 18.
FT DOMAIN 3189 3280 IG-LIKE C2-TYPE 19.
FT DOMAIN 3286 3376 IG-LIKE C2-TYPE 20.
FT DOMAIN 3384 3469 IG-LIKE C2-TYPE 21.
FT DOMAIN 3482 3572 IG-LIKE C2-TYPE 22.
FT DOMAIN 3580 3667 IG-LIKE C2-TYPE 23.
FT DOMAIN 3686 3777 IG-LIKE C2-TYPE 24.
FT DOMAIN 3817 3908 IG-LIKE C2-TYPE 25.
FT DOMAIN 3920 4009 IG-LIKE C2-TYPE 26.
FT DOMAIN 4018 4106 IG-LIKE C2-TYPE 27.
FT DOMAIN 4109 4201 IG-LIKE C2-TYPE 28.
FT DOMAIN 4212 4297 IG-LIKE C2-TYPE 29.
FT DOMAIN 4302 4387 IG-LIKE C2-TYPE 30.
FT DOMAIN 4400 4485 IG-LIKE C2-TYPE 31.
FT DOMAIN 4489 4580 IG-LIKE C2-TYPE 32.
FT DOMAIN 4588 4678 IG-LIKE C2-TYPE 33.
FT DOMAIN 4681 4771 IG-LIKE C2-TYPE 34.
FT DOMAIN 4873 4961 IG-LIKE C2-TYPE 35.
FT DOMAIN 4965 5057 IG-LIKE C2-TYPE 36.
FT DOMAIN 5067 5160 IG-LIKE C2-TYPE 37.
FT DOMAIN 5171 5260 IG-LIKE C2-TYPE 38.
FT DOMAIN 5277 5366 IG-LIKE C2-TYPE 39.
FT DOMAIN 5383 5472 IG-LIKE C2-TYPE 40.
FT DOMAIN 5487 5578 IG-LIKE C2-TYPE 41.
FT DOMAIN 5595 5685 IG-LIKE C2-TYPE 42.
FT DOMAIN 5701 5790 IG-LIKE C2-TYPE 43.
FT DOMAIN 5815 5904 IG-LIKE C2-TYPE 44.

FT DOMAIN 5925 6014 IG-LIKE C2-TYPE 45.
FT DOMAIN 6038 6130 IG-LIKE C2-TYPE 46.
FT DOMAIN 6150 6239 IG-LIKE C2-TYPE 47.
FT DOMAIN 6275 6368 FIBRONECTIN TYPE-III.
FT DOMAIN 6413 6502 IG-LIKE C2-TYPE 48.
FT DOMAIN 6507 6596 IG-LIKE C2-TYPE 49.
FT DISULFID 568 621 POTENTIAL.
FT DISULFID 2908 2975 POTENTIAL.
FT DISULFID 3015 3065 POTENTIAL.
FT DISULFID 3707 3759 POTENTIAL.
FT DISULFID 3826 3890 POTENTIAL.
FT DISULFID 5092 5157 POTENTIAL.
FT DISULFID 5298 5350 POTENTIAL.
FT DISULFID 5508 5560 POTENTIAL.
FT DISULFID 5616 5669 POTENTIAL.
FT DISULFID 5722 5764 POTENTIAL.
FT DISULFID 5836 5901 POTENTIAL.
FT DISULFID 5946 5998 POTENTIAL.
FT DISULFID 6036 6171 POTENTIAL.
FT DISULFID 6421 6486 POTENTIAL.
FT CONFLICT 2137 2137 A -> P (IN REF. 1).
FT CONFLICT 2245 2247 AKA -> PKP (IN REF. 1).
FT CONFLICT 2258 2258 A -> P (IN REF. 1).
FT CONFLICT 2284 2284 E -> G (IN REF. 1).
FT CONFLICT 2297 2297 M -> I (IN REF. 1).
FT CONFLICT 3531 3531 A -> G (IN REF. 1).
FT CONFLICT 3884 3888 DAGET -> KRRRI (IN REF. 1).
FT CONFLICT 3929 3929 A -> V (IN REF. 1).
FT CONFLICT 5145 5145 A -> P (IN REF. 1).
FT CONFLICT 5146 5145 T -> S (IN REF. 1).
FT CONFLICT 5185 5185 G -> A (IN REF. 1).
FT CONFLICT 5185 5185 K -> N (IN REF. 1).
FT CONFLICT 5199 5199 L -> F (IN REF. 1).
FT CONFLICT 5202 5202 F -> L (IN REF. 1).
FT CONFLICT 5213 5213 F -> G (IN REF. 1).
FT CONFLICT 6178 6178 A -> L (IN REF. 1).
FT CONFLICT 6268 6268 K -> E (IN REF. 1).
SQ SEQUENCE 6632 AA; 731665 MW; 262D3EDD62960B89 CRC64;

Query Match 4.7%; Score 110; DB 1; Length 6632;

Best Local Similarity 18.7%; Pred. No. 83;
Matches 97; Conservative 60; Mismatches 132; Indels 230; Gaps 25;

QY 32 KKGDVELTLCASQKSIQPHM-KNSNOIKILNGSGFLTKGPSKLNDRADRSRLMVG 90
DB 2899 KKGSHAVEECVDPDTGVCCMKLDGKEILARIIVQTRGP-----EG 2943
QY 91 NF-PLIIKNLKIEDSDTYICEVEDQKEVOLIVFGLTANSDFHLQGSITLT-LESPP 147
DB 2944 HITQELVLDNTPEDBAGKTYLIVEN-----TAGQDT-----CENTLVIESLE 2986
QY 148 GSSPS-----VQCR--SPRGK-----NIQG 165
DB 2987 KKESEKAEPIVALQDKTTKTESEKVLCEKVIQGEPRKPVSMHNKTTIQESITVESVEG 3046
QY 166 GKITSVSLQELQDGSFTWCYVLQNGKKEFFIDIVLAFOKASSIVYKKEBEQVBFPL 225
DB 3047 VERVITTSLSHQKCYTC-----IAENTESTKTEAF-- 3079
QY 226 AFTVEKLGSSELMWQAEARASSKSMITFDLKNKEVSRYKVTQPKLQMGKLLPLHLTL- 284
DB 3080 -----LTVQGE-----APVTKELQNKELST-----GKLVJLSCSVK 3111
QY 285 --PQALPOYAGSGNLT-----LALDAKTKLHQEVN---LVVMRATQ----- 321
DB 3112 GSPQGHVDFVFSFSETTKVETKITSSRAIR-----HDQNTTHMRWVISQITKEDIIVSY 3165
QY 322 -----LQKNLTCEVWGP-----TSPKMLSLKLNK-----EAKVSRRE 355
DB 3166 KAIATNSIGTATSTSKITTKVEAPVFEQGLKTKTSKEKEIKMEVKVGGASAPDVEWFFQD 3225
QY 356 KPV-----VWLNPEAGMW-----QCLDSQGVLLIESNITKVLPTWSTPVPAPASALPA 403
DB 3226 KPVSDGNHKKKMPDETGVFLVVKQATLTDAKGYTAKA-----SNPAGTJASSSAA 3277

